Physician Learning and Treatment Choices: Evidence from Brain Aneurysms

Qing Gong*

May 31, 2018

Abstract

Physicians often choose among alternative treatment options based on their beliefs over the treatment effectiveness and their skills in delivering the treatment. I examine how two kinds of physician learning jointly shape their treatment choices: *Bayesian learning* that updates beliefs about treatment-patient match values and *learning by doing* that improves surgical skills. Using case-level data on the history of brain aneurysm treatments by over 200 physicians, I find that both kinds of learning are present and that physicians are forward-looking. In light of these empirical patterns, I develop and estimate a dynamic structural model of physician learning and treatment choices for heterogeneous patients. I then quantify the impacts of the two kinds of learning and find that (a) learning encourages physicians to deviate substantially from the myopic best choices, hurting short-term patient outcomes but improving the overall treatment success rates by 13-17%; (b) learning explains 20% of total variation in physicians' choice of care, with Bayesian learning helping to reduce the variation while learning by doing adding to the variation. I also evaluate the impacts of several alternative payment schedules. Uniform payments across treatments facilitate the adoption of the new treatment while outcome-contingent payments have heterogeneous effects across physicians. The heterogeneity highlights the coexistence of two opposing effects: the incentive to exploit the myopic best option and the incentive to experiment with less familiar options due to the increased return from learning.

Keywords: Bayesian learning, learning by doing, discrete choice

*University of North Carolina at Chapel Hill. E-mail: qinggong.pku@gmail.com. I am indebted to my advisors, Hanming Fang, Ashley Swanson, and Petra Todd, for their continuing guidance and support throughout this project. I have benefited greatly from helpful comments from Juan Pablo Atal, Camilo García-Jimeno, Ami Ko, Holger Sieg, Andrew Postlewaite, and participants at the 2017 iHEA World Congress, the UPenn Empirical Micro Lunch Seminar, and the Empirical Micro Workshop. I also thank Dr. David Kung and Dr. Eric Zager at the Hospital of the University of Pennsylvania for their help with the medical and institutional backgrounds of brain aneurysm treatments. All errors are my own. I gratefully acknowledge the financial support of the Wharton Risk Management Center's Russell Ackoff Doctoral Student Fellowship in obtaining the data needed for the empirical analysis.

1 Introduction

Physicians often need to choose among alternative treatment options and are constantly learning through experience which treatments are the best match for different types of patients. They experiment with the options and update their beliefs based on patient outcomes, essentially doing *Bayesian learning*.[1] At the same time, physicians accumulate the skills to deliver each treatment via *learning by doing*. The coexistence of Bayesian learning and learning by doing are especially relevant for surgical care. Physicians' beliefs about treatment-patient matches and surgical skills affect the treatment choices and outcomes, which in turn change their future beliefs and skills.

In this paper, I study how these two kinds of learning shape physician decision-making, using the treatment of brain aneurysms as an example. I focus on how belief updating via Bayesian learning and skill accumulation via learning by doing jointly affect the treatment choices. I then evaluate the impacts of payment reforms on physician learning and patient outcomes.

The treatment of brain aneurysms is a particularly desirable setting to study the two kinds of physician learning. First, a brain aneurysm is a neurological condition with three treatment options: surgical *clipping*, the traditional option; endovascular *coiling*, a new option; *no intervention*, the outside option. Clipping and coiling are both surgical procedures, hence physician beliefs *and* skills are relevant to treatment choices. Second, the emergence of coiling as a promising yet under-explored alternative gives physicians strong incentives to learn. Third, medical guidelines for brain aneurysm treatment choices are still lacking, which further necessitates learning by individual physicians. Finally, the learning environment is fairly clean, with a limited number of options, unilateral decision-making,[2] and an immediately observable outcome measures, namely whether the patient can be discharged home without the need for assisted care.[3]

My empirical analysis is based on the New York Statewide Inpatient Database (SID). The SID covers the universe of inpatient care within the state and provides detailed case-level information on diagnoses, treatments, and outcomes. Most importantly, it allows me to track physicians across hospitals and years, thereby retrieving the uninterrupted history of brain aneurysm cases by physician.

---

[1]See, for example, Coscelli and Shum (2004) and Crawford and Shum (2005) on physician learning about anti-ulcer drug choices; Dickstein (2014) on anti-depressant drugs; and Saxell (2014) on cholesterol drugs.

[2]The treatment of brain aneurysms features strong information asymmetries between physicians and patients and barely any shared decision-making. See The Dartmouth Atlas Working Group (2015) for a detailed discussion.

[3]This binary outcome measure is widely used in the medical literature and has been shown to be a good proxy for patients' longer term health outcomes (Zacharia et al., 2014).

I start with reduced-form evidence of whether and how physicians learn. First, I show that a physician is more likely to choose a treatment if she has used it on more patients or has had better patient outcomes with it on *similar patients*. The path dependence suggests the presence of skill accumulation and patient type-specific belief updating, respectively. Using the subsample of emergency cases, I show that the dependence is not driven by the sorting of patients. I also rule out learning from peers by showing that a physician's utilization rate of a treatment is not affected by that of her colleagues, once conditional on her own history. Second, I find that the average patient outcome improves over time, but such improvement is mainly driven by better treatment-patient matches instead of skill accumulation. The outcomes of similar patients no longer vary over time, once conditional on receiving the same treatment. That is, while both learning by doing and learning about treatment-patient match values affect physician *choices*, they have different effects on patient *outcomes*. Third, I document suggestive evidence that physicians are forward-looking.[4] Physicians with high future patient arrival rates are more likely to adopt coiling, the new option, in the earlier periods. The response to future patient arrivals suggests physicians are likely to be forward-looking, because a myopic physician would not account for the return of learning, which partially depends on the frequency of patient arrivals.

The reduced-form results motivate a model with forward-looking physicians who mainly learn from their own experiences, accumulating skills and updating beliefs about treatment-patient matches. The choice dynamics and the intertemporal tradeoff in learning also necessitate a structural model that can disentangle the two kinds of learning and evaluate alternative policies that change physician learning incentives.

I develop a dynamic model of physician treatment choices under two distinct kinds of learning. Each physician holds (a) a set of beliefs about the latent and invariant match value between each pair of treatment and patient type and (b) a set of evolving, treatment-specific surgical skills. The forward-looking physician then makes treatment choices for a sequence of heterogeneous patients, maximizing the discounted sum of her expected payoffs. The physician's flow payoff consists of the expected patient outcome, which depends on her beliefs regarding treatment-patient match values; the cost of delivering a treatment, which decreases as she accumulates more surgical skills in that treatment; and the expected treatment revenue, which reflects her financial incentives. Finally, the physician's beliefs and skills evolve after treating each patient. The physician starts with hetero-

[4]The tests follow Abaluck et al. (2015), who identifies the extent to which Medicare Part D enrollees are forward-looking from their response to *future* prices.

geneous prior beliefs about the match value between each pair of treatment and patient type. She updates her belief in a Bayesian fashion after observing the outcome of the particular treatment-type combination. At the same time, her surgical skill in that treatment grows deterministically regardless of the patient type or the outcome.

The model features a high-dimensional state space and spillovers between the two kinds of learning; that is, an increase in the skill of one treatment affects the physician's future evaluation of that treatment for *all* types of patients. To address these challenges, I solve the model by modifying the Gittins index, which uses forward induction to circumvent the curse of dimensionality (Gittins, 1979). The modified Gittins index accommodates the spillover of learning by doing across patient types, provides a sufficient statistic for the value of each option, and only depends on the current state.[5] I follow Whittle (1982) and prove that always choosing the option with the highest index is an optimal policy.

I estimate the model on the SID data using maximum likelihood. The learning parameters imply that physician beliefs converge after Bayesian learning from 10-20 cases, and that the accumulation of skills has large impacts on treatment choices: a physician will be indifferent between an option she has used twice and another option she believes to have a 6.5-percentage-point lower success rate but has used 7 times, the average annual caseload in the sample. Using simulation, I disentangle the impacts of the two kinds of learning. The effect of Bayesian learning dominates that of learning by doing on the utilization of clipping, explaining 78% of the changes in clipping probability when both kinds of learning are shut down. This reflects the high initial stock of skills in clipping, which limits further learning by doing. On the contrary, the effect of Bayesian learning is dominated by that of learning by doing on the adoption of coiling, explaining only 26% of the changes.

With counterfactual experiments, I first quantify the effects of learning in two ways. I begin by exploring how physicians would choose differently if they were myopic and had disregarded the value of learning. I find that the adoption rate of coiling would decrease from 41% to 31% if physicians were myopic. In particular, forward-looking physicians are more than twice as likely to deviate from the myopic best choices and experiment with coiling on unhealthy patients than healthier ones.[6] I also find that the experimentation hurts short-term outcomes, but the overall treatment success rates are 13-17% higher than they would be with myopic physicians. I then

[5]As pointed out by Dickstein (2014), the Gittins index also provides an intuitive rule of thumb for physician decision-making, which may be more relevant in practice.

[6]This echoes the intuition of the "stepping stone" model by Jovanovic and Nyarko (1997), in which agents first experiment with options that suffer from smaller losses in case of a failure.

gauge the contribution of learning to the variation in physicians' choice of care. I find that shutting down both channels of learning would reduce total variation by 20%. [7] In addition, the two kinds of learning work differently: Bayesian learning helps to reduce the variation through belief convergence; learning by doing adds to the variation by solidifying the randomness in physicians' past outcomes, thereby driving physicians down diverging paths even if they were similar ex ante.

In another set of counterfactuals, I evaluate the impacts of two payment reforms. I find that uniform payments across treatments facilitate the adoption of coiling for almost all physicians, as current choices are partly influenced by the lower payment for coiling. I also examine outcome-contingent payments such as the on-going Value Modifier (VM) program by the Centers for Medicare and Medicaid Services. VM pays physicians 102% of the base amount for good patient outcomes and 98% for bad ones. I find that the responses to VM are heterogeneous across physicians, highlighting the classic *exploitation versus experimentation* tradeoff in learning. The increased reward from a success induces physicians to exploit the myopic best option while the greater return of learning encourages experimentation with less familiar options.

My study contributes to three strands of literature. First, I build on the literature of physician learning by explicitly incorporating physician skill accumulation into a Bayesian learning model. Previous studies have primarily focused on Bayesian learning about prescription drugs (e.g., Coscelli and Shum (2004); Crawford and Shum (2005); Ferreyra and Kosenok (2011)).[8] Dickstein (2014) examines physician Bayesian learning in a multi-armed bandit framework with forward-looking physicians and correlated learning across similar drugs. My study is the first to model physician skill accumulation in tandem with belief updating. The inclusion of the learning-by-doing channel is especially relevant in the context of surgical conditions such as brain aneurysms. I complement previous studies by allowing for endogenous beliefs *and* skills, which jointly shape physicians' treatment choices and evolve depending on those choices. Moreover, I am able to estimate an otherwise high-dimensional model with the modified Gittins index, which accommodates forward-looking physicians and spillovers across the two kinds of learning.

Second, my work complements an earlier literature on worker learning in the labor market. Jovanovic (1979) proposes a model in which the worker learns about his productivity on each job.

[7]Previous studies (e.g. Finkelstein et al. (2016)) find that the supply side (healthcare providers) accounts for about 60% of total variation. Hence learning alone contributes a significant fraction of the supply-side variation.

[8]Also see Ching et al. (2013) for an excellent survey of studies on consumer learning in experience good markets. Other studies extend the focus beyond physicians' learning from their own experiences and examine learning from watching peers (Ho, 2002), detailing (Narayanan and Manchanda, 2009), patients' prescription history (Saxell, 2014), and public disclosure of physician performances (Kolstad, 2013).

Jovanovic and Nyarko (1997) focus on worker mobility and compare the stepping-stone model, in which worker productivities evolve on the job, with the bandit model à la Miller (1984), in which workers learn about their latent match values with each job. Recent work by Papageorgiou (2014) studies workers who learn about their comparative advantages in different occupations. My paper focuses on a combination set of learning objectives that are more relevant in the context of medical decision-making: physicians try to recover the *latent* match value between treatments and patients via Bayesian learning while simultaneously accumulating surgical skills in each treatment via learning by doing.

Finally, I contribute new empirical evidence to the large literature on physician practice styles and variations in the choice of care.[9] On the demand side, studies find that patient preferences are relatively unimportant (Cutler et al., 2013; Baker et al., 2014; Finkelstein et al., 2016). On the supply side, researchers have looked into how physician behaviors are impacted by factors such as financial incentives (Gruber et al., 1999; Johnson and Rehavi, 2016), litigation risks (Baicker et al., 2007), team composition (Chan, 2016), and practice environments (Molitor, 2018). Even for completely benevolent physicians, specialization can also affect treatment decisions by changing the return of different treatment options (Chandra and Staiger, 2007). Other studies point to more tacit characteristics of physicians. The heterogeneity in physician beliefs (Cutler et al., 2013), aggressiveness (Abaluck et al., 2016), procedural skills and the responsiveness to patient conditions (Currie et al., 2016; Currie and MacLeod, 2017) all have significant impacts on the choice of care and patient outcomes. I add to this literature by documenting the within-physician choice *dynamics* with a novel dataset that includes physician histories and observable outcomes, which are rarely available in previous studies.

The paper proceeds as follows. In Section 2, I introduce the background on brain aneurysm treatments, describe the SID data, and document reduced-form evidence on physician learning. I then develop a dynamic model of physician learning and treatment choices in Section 3. I specify the econometric model and discuss the identification strategy in Section 4. In Section 5, I present the estimation results and disentangle the impacts of the two kinds of learning. With counterfactual experiments in Section 6, I show how physician myopia and payment reforms would influence treatment choices and outcomes. I conclude in Section 7.

[9] See Skinner (2011) and Chandra et al. (2011) for comprehensive surveys of related studies in both the economic and the clinical literature.

2 Empirical Backgrounds and Data

2.1 Brain aneurysms and treatment options

Brain aneurysms (also known as *cerebral aneurysms* or *intracranial aneurysms*) are blisters formed on weakened spots of brain arteries. Vlak et al. (2011) estimated that about 3.2 percent of the population have brain aneurysms. The average age of brain aneurysm patients is 50, with a large variance. Despite their prevalence, the detection of brain aneurysms is not as common because the majority of aneurysms are asymptomatic. Diagnoses have only increased starting in the early 2000s with the advances of imaging technologies (Wiebers et al., 2003).

Although mostly asymptomatic, brain aneurysms can be life-threatening if they rupture, which happens with an average probability of 1% per year (Wiebers et al., 2003). Ruptured aneurysms will cause bleeding inside the brain (*subarachnoid hemorrhage*, SAH). 10-15% of SAH sufferers die before reaching the hospital. Even for those who manage to reach the hospital and receive emergency care, the mortality rates are still as high as 25%.

There are three options for managing brain aneurysms. The first option, surgical *clipping*, is the traditional treatment for brain aneurysms since the 1920s. The neurosurgeon cuts an opening in the skull and places a clip across the neck of the aneurysm to stop the blood flow into the aneurysm.[10] This treatment is both durable and effective in preventing ruptures, yet is invasive and poses relatively high mortality and morbidity risks. The second option, endovascular *coiling*, is a more recent procedure approved by the Food and Drug Administration (FDA) in 1997. Its take-up rates were below 10 percent until a 2003 study in *The Lancet* showing its comparable outcomes to clipping (Wiebers et al., 2003). But the medical community is still uncertain about how well it suits different types of patients. The third, outside option is watchful *observation* and no intervention. In this case, the neurosurgeon orders follow-up diagnostic imaging tests every 6-12 months to monitor the development of the aneurysm.

A commonly used measure of treatment success is whether the patient can be discharged home and does not need assisted care (Zacharia et al., 2014). This binary indicator is shown to be a good predictor of the patient's long-term health and wellbeing. Moreover, it is clearly defined and observable to the attending physician. Hence the physician can easily get feedback on her treatment, thereby learning from it. It is also readily available in hospital records, so I am able to observe the treatment outcomes in the inpatient care data that I use.

[10] Figure A1 in the Appendix illustrates the procedure.

Table 1 summarizes the key differences among the three treatment options using data from New York State from 2003-2014. Both options of intervention are costly and require several days of inpatient stay. The median total charge of clipping is 36% more expensive than coiling, partially driven by the extra cost incurred during the significantly longer stay. The bottom half of the table shows the different match values between treatments and patient types. The probability of having treatment success rates is higher for patients with unruptured aneurysms, as well as patients with no major comorbidities.

	Surgical clipping	Endovascular coiling	Watchful observation
Since	1920s	1997	-
Invasiveness	Invasive craniotomy	Minimally invasive	-
Median charges ($1,000)	123.72	90.34	7.55
Median inpatient days	12	4	2
Treatment success rates by patient type:			
Ruptured, no major comorbidity	27.1%	40.0%	37.7%
Ruptured, with major comorbidity	29.0%	35.1%	16.6%
Unruptured, no major comorbidity	73.4%	93.1%	90.7%
Unruptured, with major comorbidity	59.1%	84.3%	84.9%

Table 1: Treatment options for brain aneurysms
NOTES: Charges are in real 2014 thousands of dollars and include all charges for the inpatient stay. A treatment is a success if the patient can be discharged home and does not need assisted care. Ruptured and unruptured aneurysms are identified from standardized diagnosis codes (ICD-9-CM) in the SID data. Healthy patients are those with no major comorbidities (diseases or disorders in addition to the primary condition, i.e. brain aneurysm) as recorded in the SID.

2.2 The New York Statewide Inpatient Database (2003-2014)

I use the State Inpatient Database (SID) for New York from 2003 to 2014 for my empirical analyses. The SID includes the *universe* of inpatient discharges at almost all of the over 200 hospitals in New York State. More importantly, each physician has a unique ID that stays unchanged across hospitals and time. Hence I am able to identify all the brain aneurysm cases treated by a given physician over the sample period, as long as the procedures are done within the state borders.[11] This is crucial for retrieving the entire history of a physician's treatment choices and patient outcomes over time.

Each observation in SID is an inpatient discharge and has a rich set of information on the timing,

[11] Due to licensing requirements, practicing neurosurgery in multiple states is not common. Thus I will assume the neurosurgeons' brain aneurysm cases in the State of New York capture *all* of their inpatient brain aneurysm cases.

source, and source of the admission; demographic and medical information of the patient; up to 15 standardized codes on diagnoses and treatments; the cost, length, and, most importantly, the outcome of the inpatient stay.

I construct the main sample in two steps. First, I impose the following selection criteria on the SID: I exclude maternal or neonatal admissions (20.09%) to focus inpatient admissions with disease-related medical needs. I also exclude admissions that are transferred from law enforcement (0.08%), those with missing physician identifiers (0.47%), or those with the patient's age under 15 (15.31%). Admissions that belong to *one* of the above cases account for 25.78% of the whole SID panel and are dropped in this step. The remaining sample has 22.40 million inpatient cases. Second, I identify and focus on the subsample of admissions due to brain aneurysms using the standardized ICD-9-CM diagnosis codes following Brinjikji et al. (2011). I further identify and exclude emergency cases with only first-aid procedures and resulted in patient deaths, which indicate limited room for medical decision-making.

The resulting main sample includes 11,767 brain aneurysm cases (10,629 unique patients) treated by 219 physicians at 111 hospitals. Note that the vast majority of patient-physician encounters are only one-shot. About 50% of the cases are emergency ones where patients can barely search for or be referred to certain physicians. This subsample will help me address the potential endogeneity issues in Section 2.3. Table 2 summarizes the workload, choice patterns, and professional affiliations of the physicians in the sample.

2.3 Reduced-form evidence

I draw upon the SID data for empirical patterns of treatment choices. I explore (a) whether Bayesian learning and learning by doing are present, (b) how the two kinds of learning affect patient outcomes differently if they are present, and (c) to what extent physicians are forward-looking.

Presence of both kinds of learning. My first set of reduced-form results provide evidence that both kinds of learning are present and affect physicians' treatment choices. Figure 1 illustrates the convergence in physician choices over time. *Panel A* plots the cumulative shares of patients treated by clipping (left) and coiling (right) by physicians who have high clipping adoption rates when they are last observed in the SID data (i.e. in 2014 or when they leave the sample, whichever comes earlier). The share of cases treated by these physicians with clipping is dispersed in the earlier months, but converges to a relatively high level over time. The share of cases treated with coiling

	Mean	S.D.	Min	Max
A. Physician workload				
Annual caseload	7.48	9.16	1	67
Annual number of patients	6.98	8.42	1	67
Annual number of cases treated with clipping	2.25	4.27	0	37
Annual number of cases treated with coiling	3.07	5.92	0	48
B. Physician choice patterns				
Share of cases treated with clipping	0.30	0.46	0	1
Share of cases treated with coiling	0.42	0.49	0	1
Share of cases treated with observation	0.29	0.45	0	1
Share of cases treated with clipping, conditional on intervention	0.42	0.49	0	1
Share of cases treated with coiling, conditional on intervention	0.58	0.49	0	1
C. Physician experience and professional affiliations				
Fraction of young physicians	0.35	0.48		
Fraction of teaching-hospital physicians	0.33	0.47		
Number of hospitals worked at	1.46	0.96	1	8
Number of physicians at the same hospital in a month	1.86	1.16	1	9
Number of physicians at the same hospital in a year	3.25	2.55	1	13

Table 2: Physician workload, choice patterns, and professional affiliations in the main sample

NOTES: The main sample is constructed from the New York SID (2003-2014) and includes 11,767 inpatient cases for which aneurysms are the primary cause of admission. *Caseload* refers to the number of inpatient cases treated by a neurosurgeon in a year. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.*Number of hospitals worked at* is the number of hospitals a neurosurgeon every practiced in. *Number of physicians at the same hospital in a month (year)* is the number of other neurosurgeons who treat brain aneurysms at the same hospital as a physician in a calendar month (year).

is also dispersed at the beginning, but converges to 0-20%. *Panel B* shows similar patterns among physicians who end up with high coiling adoption rates.

I then show that physician choices respond to past experiences and outcomes with the treatments. Table 3 reports the estimates from multinomial logit regressions of the case-level choice probabilities. I regress physician i's choices for the patient in period t on the physician's past experiences with each treatment and the resulting patient outcomes. Columns (1) and (2) show that physicians are more likely to choose a treatment if they have used it more on previous patients; they also tend to favor the treatment with which they have had better outcomes *on similar patients*. The path dependence suggests that both skill accumulation (learning by doing) and patient type-specific belief updating (Bayesian learning) are present. In Columns (3) and (4), I use the subsample of emergency cases where there is limited room for selection or sorting based on patient characteristics. The choice patterns are highly similar to those using the whole sample, thereby ruling out the endogeneity problem.

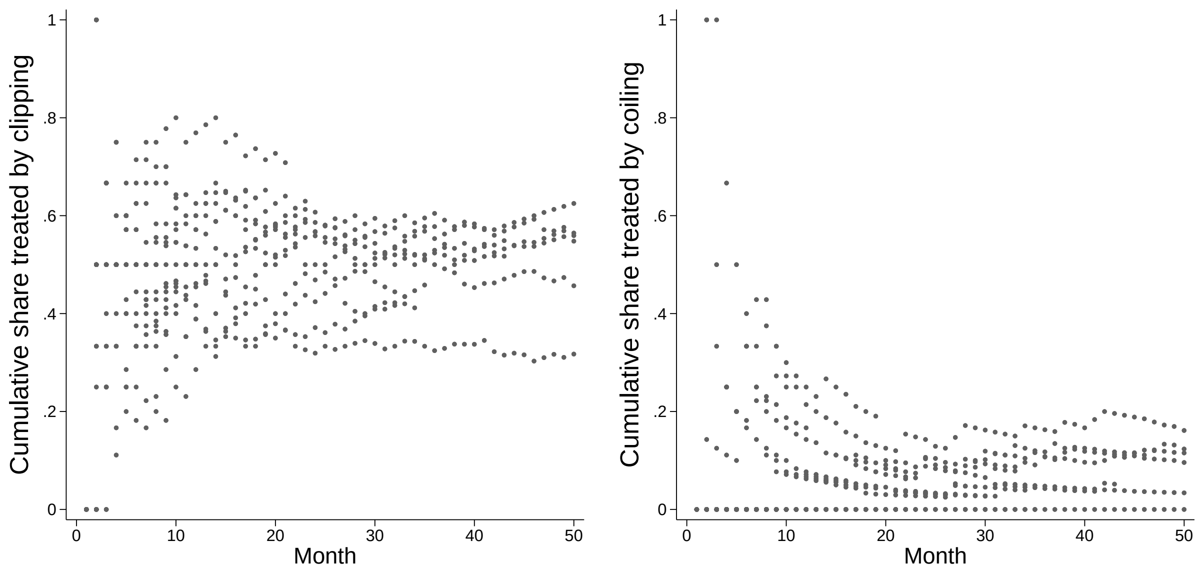
Dependent variable: case-level treatment choice $d \in \{Clip, Coil, Obs\}$				
	Whole sample		Emergency cases	
	Pr(clipping)	Pr(coiling)	Pr(clipping)	Pr(coiling)
	(1)	(2)	(3)	(4)
ln(previous clipping cases)	0.068***	-0.044***	0.077***	-0.050***
	[0.004]	[0.004]	[0.006]	[0.006]
ln(previous coiling cases)	-0.046***	0.048***	-0.043***	0.071***
	[0.005]	[0.005]	[0.007]	[0.006]
ln(previous clipping successes*)	0.147***	-0.057***	0.137***	-0.023**
	[0.006]	[0.007]	[0.009]	[0.009]
ln(previous coiling successes*)	-0.077***	0.162***	-0.048***	0.128***
	[0.006]	[0.005]	[0.011]	[0.009]
Observations	11,746		6,296	

Table 3: Reduced-form evidence: treatment choices respond to past experiences and outcomes

NOTES: The table reports the marginal effects on choice probability from a multinomial logit regression, where the dependent variable is physician i's choice (clipping, coiling, or observation) for the patient in period t. *ln(previous cases of clipping)* is the log number of cases physician i has treated with clipping up to but not including t. *ln(previous clipping success*)* is the cumulative number of physician i's cases treated with clipping where (i) patients are of the same type as the current patient and (ii) the outcome was a success. A treatment is a success if the patient can be discharged home and does not need assisted care. Columns (1)-(2) use the entire sample; (3)-(4) use the subsample of emergency cases to rule out the alternative explanation that the path dependence shown in (1)-(2) is driven by the sorting of patients over time. Other covariates include patient demographics, insurance status, comorbidities, aneurysm types, sources of admission, hospital fixed-effects, and year fixed-effects. Standard errors are reported in brackets.** $p < 0.05$, *** $p < 0.01$

I further rule out a physician's peers as the major source of learning. Previous studies have found the impact of team learning or learning by watching on physician choices (Ho, 2002; Reagans et al., 2005). But in the SID sample, I observe only one physician treating brain aneurysms

Panel A. Convergence of choices by high-clipping physicians



Panel B. Convergence of choices by high-coiling physicians

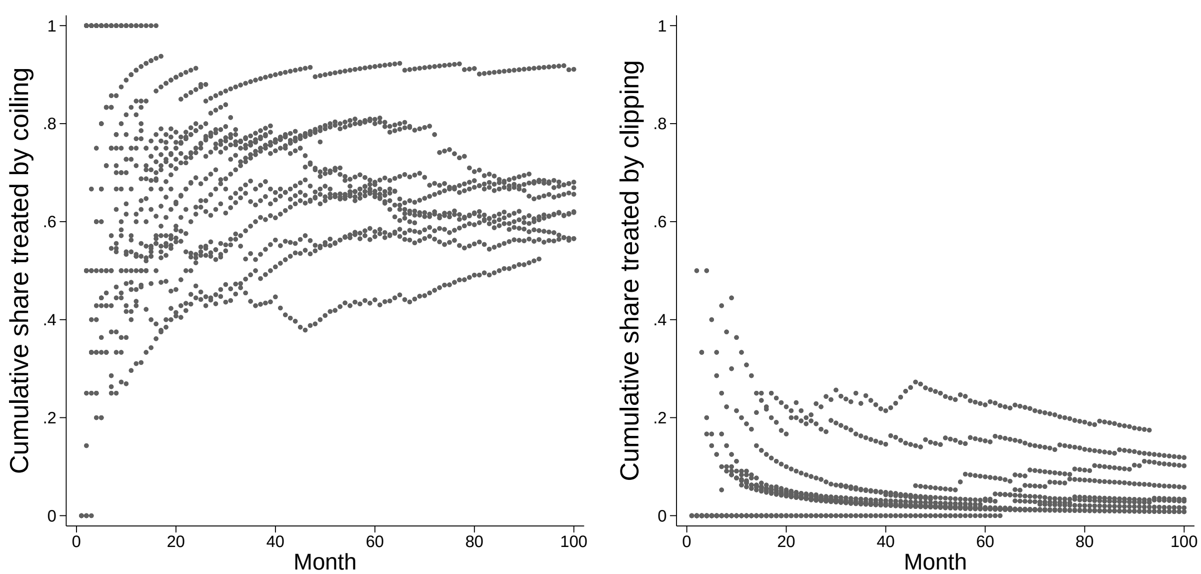


Figure 1: Convergence of choices over time

Notes: *Panel A* plots the cumulative shares of patients treated by clipping (left) and coiling (right) by physicians who have high clipping adoption rates when they are last observed in the SID data (i.e. in 2014 or when they leave the sample, whichever comes earlier). The share of cases treated by these physicians with clipping is dispersed in the earlier months, but converges to a relatively high level over time. The share of cases treated with coiling is also dispersed at the beginning, but converges to 0-20%. *Panel B* plots the same convergence patterns in the shares of coiling (left) and clipping (right) by physicians who end up with high coiling adoption rates.

at each hospital more than 50% of the time. In addition, I show in Table 4 that a physician's utilization of clipping and coiling respond to her own but not her peers' choices. Columns (1)-(2) use the subsample of physicians who have at least one colleague at the same hospital within a calendar year. I find that while the physician's utilization is strongly affected by her past choices and outcomes, it is not correlated with her peers' choices. The physician's utilization rate of coiling is even negatively correlated with her peers', although the magnitudes are tiny compared with other factors. Columns (3)-(4) use the smaller subsample of physicians who have colleagues at the same hospital within a calendar *month*, and find similar results.

Dependent variable: physician i's utilization rate of clipping or coiling				
	Clip%	Coil%	Clip%	Coil%
	(1)	(2)	(3)	(4)
Own share of clipping		-0.5722***		-0.5869***
		[0.0116]		[0.0130]
Own share of coiling	-0.3963***		-0.3920***	
	[0.0081]		[0.0087]	
Own success rate with clipping	0.0950***	0.1606***	0.0967***	0.1808***
	[0.0058]	[0.0069]	[0.0064]	[0.0077]
Own success rate with coiling	-0.0411***	0.2944***	-0.0400***	0.3037***
	[0.0064]	[0.0070]	[0.0071]	[0.0079]
Peers' share of clipping	-0.0011	0.0042***	-0.0012	0.0047***
	[0.0009]	[0.0011]	[0.0012]	[0.0014]
Peers' share of coiling	-0.0012	-0.0069***	-0.0009	-0.0071***
	[0.0009]	[0.0011]	[0.0010]	[0.0013]
Adjusted R^2	0.5387	0.6301	0.5116	0.6197
Observations	8,402	8,402	6,870	6,870

Table 4: Reduced-form evidence: physicians are not following peers' choices

Notes: The table reports estimates from linear regressions with the dependent variable being a physician's cumulative shares of clipping ((1) and (3)) and coiling ((2) and (4)). Columns (1)-(2) use the subsample of physicians who have at least one colleague at the same hospital within a calendar year. Columns (3)-(4) uses the subsample of physicians who have at least one colleague at the same hospital within a calendar month. A treatment is a success if the patient can be discharged home and does not need assisted care. Other covariates include year and hospital fixed effects. Standard errors are reported in brackets. *** $p < 0.01$

Different outcome implications between Bayesian learning and learning by doing. My second set of reduced-form results look into the change in patient outcomes over time. Although previous results show that both beliefs and skills influence physician *choices*, I find that they have different implications for patient *outcomes*. Column (1) of Table 5 shows that the outcomes of a physician's patients do improve over time. But the improvement is likely the result of more accurate beliefs and better matching between treatments and patients, not skill accumulation: column (2) shows that once conditional on receiving the same treatment, the patient outcomes no longer vary discernibly over time. This will motivate my specification of the structural model, in which physician beliefs affect the expected outcome, but skills only help to reduce the physician's cost of delivering care.[12]

Dependent variable: 1(treatment success)		
	(1)	(2)
ln(cumulative number of patients)	0.105**	-0.0257
	[0.050]	[0.0316]
Physician fixed effects	Y	Y
Patient type dummies	Y	Y
Treatment dummies	N	Y
Observations	8,709	

Table 5: Reduced-form evidence: evolution of patient outcomes
NOTES: The table summarizes the results from a panel regression with physician fixed effects, where the dependent variable is whether the treatment is a success. A treatment is a success if the patient can be discharged home and does not need assisted care.*ln(previous cases of clipping)* is the log number of cases physician i has treated with clipping up to but not including t. *ln(cumulative number of patients)* is the total number of patients the physician has treated up to but not including the current patient. Other covariates include patient demographics, insurance status, comorbidities, aneurysm types, and sources of admission. Column (1) also controls for patient types. Column (2) controls for both patient types and the chosen treatment. Standard errors are reported in brackets.** $p < 0.05$

Forward-looking physicians. In the last set of reduced-form analyses, I test whether physicians' initial choices also respond to patient arrival rates in the *future*. The evidence will shed light on the extent to which physicians are forward-looking: the return from learning is higher if a physician expects to have frequent arrivals of brain aneurysm patients, or if the current patient is of a common type. Hence a forward-looking physician will be more likely to experiment with the lesser known treatment option in these scenarios, whereas a myopic physician will be indifferent. Table 6 indeed shows that physicians with more experience in the past are more likely to choose clipping than coiling in the early periods, which reflects some degree of path dependence. But experienced physicians who have higher future patient arrival rates are more likely to choose coiling over clipping than their experienced peers with low arrival rates. Table 7 also finds that a *given* physician is more likely to choose coiling on a patient if she expects more patients with the same type in the future.

The reduced-form results provide strong evidence of learning by physicians who are forward-looking, accumulate skills, and update beliefs from their own experience. The rich choice dynamics also call for a structural model to disentangle the two kinds of learning and to do counterfactual

[12]Stone and Bernstein (2007) studied 2684 errors in 1108 neurosurgical cases, and found that only 2.7% of all errors have significant impacts on patient outcome.

Dependent variable: case-level treatment choice $d \in \{Clip, Coil, Obs\}$		
	Pr(clipping)	Pr(coiling)
1(experienced physician)	0.143***	-0.141***
	[0.0303]	[0.0357]
1(Low arrival rates)×1(experienced physician)	0.153***	-0.142***
	[0.0205]	[0.0342]
1(High arrival rates)×1(experienced physician)	-0.144***	0.139***
	[0.0298]	[0.0348]
Observations	6,276	

Table 6: Reduced-form evidence: initial treatment choices respond to future patient arrival rates

Notes: The table reports the marginal effects on choice probability from a multinomial logit regression, where the dependent variable is physician i's choice (clipping, coiling, or observation) for the patient in period t. The sample is restricted to the first 15 cases of each physician in order to show how initial choices respond to future patient arrivals. Experienced physicians are those with at least 6 years of experience. *1(Low arrival rates)* indicates the physician's future monthly patient arrival rate is below 30%. *1(High arrival rates)* indicates the physician's future monthly patient arrival rate is above 70%. Other covariates include arrival rate dummies, patient characteristics, teaching hospital dummy, and year fixed-effects. Standard errors are reported in brackets.*** $p < 0.01$

Dependent variable: case-level choices and outcomes			
	(1)	(2)	(3)
	Pr(clipping)	Pr(coiling)	Pr(success)
Type arrival rate	0.0979***	0.210***	-0.195***
(%)	[0.0306]	[0.0314]	[0.0273]
Physician fixed-effects	Y	Y	Y
Observations	4,766	4,766	4,766

Table 7: Reduced-form evidence: initial treatment choices respond to future *type-specific* patient arrival rates

Notes: The table reports estimates from panel regressions with physician fixed effects, where the dependent variables are (1) the probability of choosing clipping; (2) the probability of choosing coiling (column 2); (3) the probability of treatment success. A treatment is a success if the patient can be discharged home and does not need assisted care. The sample is restricted to the first 15 cases of each physician in order to show how initial choices respond to future patient arrivals. *Type arrival rate* is the probability of having a patient with the same type as the current patient in any given month in the future. Other covariates include patient characteristics and initial market shares of treatment options in the quarter of the physician's entry. Standard errors are reported in brackets.*** $p < 0.01$

experiments. In the next section, I develop such a model in light of the empirical patterns to study how Bayesian learning and learning by doing jointly shape physicians' treatment choices.

3 A Dynamic Model of Physician Learning

I first develop a dynamic model featuring forward-looking physicians who make treatment choices for patients with different observable types. I then show how physician beliefs and skills evolve in two separate learning processes. The model highlights the tradeoff between *exploiting* high flow payoffs and *exploring* lesser-known options for information and skills that may generate high continuation values. I solve the model by modifying the standard Gittins index (Gittins, 1979; Whittle, 1982) to circumvent the curse of dimensionality and to accommodate the spillover of learning by doing across patient types. I conclude the section by characterizing the modified Gittins index following the numerical approximation by Brezzi and Lai (2002).

3.1 Model setup

The decision maker and time. Consider the decision-making process of physician i treating patients with a given condition over an infinite planning horizon, $t = 0, 1, 2, ...$ Each time period is a month. Physicians share a common discount factor, $\tilde{\beta} \in (0, 1)$.

Patients. At most one patient arrives in each period, with probability λ_{i0} no patient arrives. Hence I index the patient arriving in period t by t for notational simplicity. Each patient has an observable type, $k_t \in \mathcal{K}$, where $\mathcal{K} = \{1, \ldots, K\}$ is a finite set.[13] The distribution of k_t is i.i.d. across time and known to the physician. Let λ_{ik} be the probability that patient t is of type k conditional on arrival, subject to the constraint $\sum_{k\in\mathcal{K}} \lambda_{ik} = 1$.

Treatment options. The physician is to choose a treatment for each patient t. Denote the set of available treatments by $\mathcal{D}$, which comprises three options: watchful observation ($d = 0$), surgical clipping ($d = 1$), and endovascular coiling ($d = 2$). Let θ^{dk} be the match value between option d for type-k patients, which governs the distribution of patient outcome y, $F(y; \theta^{dk})$. The true match values are time-invariant but unknown to the physician and will be learned over the course of a physician's practice.

[13] For convenience, I denote the case of no patient arrival by having a patient of type 0.

Physician beliefs, skills, and flow payoffs. [14] At the beginning of period t, the physician holds a complete set of beliefs and skills, $(\{\boldsymbol{\theta}_t^d\}_{d\in\mathcal{D}}, \{e_t^d\}_{d\in\mathcal{D}})$. $\boldsymbol{\theta}_t^d = (\theta_t^{d1}, \ldots, \theta_t^{dK})'$ is a $K \times 1$ vector of beliefs about the match value of treatment d for each type of patient. e_t^d is the physician's cumulative experience in d up to, but not including, period t. Note that the beliefs, $\boldsymbol{\theta}_t^d$, are type-specific but physician experience e_t^d is not—treating any type of patient with d improves the physician's surgical skills for d.

The physician's choice-specific flow utility is

$$u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk}) = \begin{cases} \mathbb{E}[y_t^{dk} \mid \theta_t^{dk}, k_t] + \alpha r_t^{dk} - c(e_t^d), & \text{if } k_t \geq 1 \\ 0, & \text{if } k_t = 0 \end{cases} \tag{1}$$

I normalize the payoff to 0 when there is no patient. When there is a patient of type $k_t = k$, the payoff has three components. First, the physician cares about the patient's expected outcome, $\mathbb{E}[y_t^d \mid \theta_t^{dk}, k_t = k]$. She forms the expectation using her belief θ_t^{dk} at the beginning of period t. Second, the physician takes into account the revenue she generates, r_t^{dk}. Because the physician does not know the exact amount at the time of decision-making, she uses the hospital average for type-k patients treated with d in the same year as t. The physician observes r_t^{dk} at the beginning of each period and holds naive expectations that future revenues remain unchanged. The expected revenues are weighted by α in the physician's utility. Third, the physician subtracts from her flow payoff the (physical or psychological) cost to deliver treatment d, $c(e_t^d)$. The cost shrinks as the physician accumulates more skill, e_t^d.

Discussion: effects of Bayesian learning vs. learning by doing. The physician improves her treatment-specific surgical skills, e_t^d, and beliefs about the treatment-patient match value, θ_t^{dk}. The two kinds of learning have different effects on future physician payoff in my model: belief updating changes the physician's expectation of outcome for a treatment-patient pair; skill accumulation lowers treatment costs for the physician. I choose this modeling approach in light of the empirical patterns discussed in Section 2.3. An alternative modeling approach is to let both the belief and the surgical skill affect the patient outcome, y^{dk}. In that case, the physician's belief updating will depend on the realized outcome *and* her skill level when furnishing the treatment.

[14] I suppress the i subscript from now on.

3.2 The learning processes

3.2.1 Learning by doing and skill accumulation

For clipping ($d = 1$) and coiling ($d = 2$), the physician's treatment-specific skills accrue as she treats more patients:

$$e^d_{t+1} = e^d_t + \mathbf{1}\{d_t = d\} \tag{2}$$

The evolution of skills is independent of patient type and treatment outcome. That is, the additional experience gained from treating a type-k patient with d has spillover effects on all types of patients to be treated with d in the future.

A physician who is more experienced with clipping or coiling incurs lower costs to deliver the treatment. The $c(\cdot)$ function maps the stock of skills, e^d_t, into the cost of doing d in period t.

I assume that the physician's experience for watchful observation ($d = 0$) is fixed at zero. The physician does not accumulate experience because she merely refers the patient to a neurologist for regular brain scans, which neither improves nor requires her own skills. Consequently, the cost of $d = 0$ is also fixed.

3.2.2 Bayesian learning and belief updating

Let Y^d_t be the latent outcome of patient t when treated with d. Assume Y^d_t follows a Bernoulli distribution with success rate θ^{dk}:

$$Y^d_t = \begin{cases} 1, & \text{with probability } \theta^{dk} \\ 0, & \text{with probability } (1 - \theta^{dk}) \end{cases} \tag{3}$$

where k is the type of patient t, and θ^{dk}_0 is the latent match value that the physician needs to learn. The physician holds Beta-distributed prior beliefs about θ^{dk}

$$\theta^{dk}_0 \sim \text{Beta}(a^{dk}_0, b^{dk}_0) \tag{4}$$

where $(a_0^{dk} > 0, b_0^{dk})$ are strictly positive parameters. They determine the mean and variance of the distribution as follows:

$$\mu_0^{dk} = \frac{a_0^{dk}}{a_0^{dk} + b_0^{dk}} \tag{5}$$

$$\nu_0^{dk} = \frac{a_0^{dk} b_0^{dk}}{(a_0^{dk} + b_0^{dk})^2 (a_0^{dk} + b_0^{dk} + 1)} \tag{6}$$

Suppose the physician has treated n_t^{dk} cases of type-k patients with d at the beginning of period t. Further suppose that s_t^{dk} cases are successful. Then the physician's posterior belief is a Beta distribution with parameters

$$a_t^{dk} = a_0^{dk} + s_t^{dk} \tag{7}$$

$$b_t^{dk} = b_0^{dk} + (n_t^{dk} - s_t^{dk}) \tag{8}$$

The posterior mean is higher if s_t^{dk} is larger, holding n_t^{dk} constant; the posterior variance will be larger if s_t^{dk} and $(n_t^{dk} - s_t^{dk})$ are close. That is, the physician is more optimistic about the match value of d for type-k when she sees more successes in the past, and is more confident in her beliefs when past successes outnumber failures by a larger margin.

3.3 Model solution: the modified Gittins index policy

3.3.1 The physician's dynamic problem

The physician maximizes the total discounted expected payoff by choosing the optimal sequence of treatments

$$\max_{\{d_t\}_{t=0,1,\ldots}} \mathbb{E}[\sum_{t=0}^{\infty} \tilde{\beta}^t u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk}) \mid \boldsymbol{\theta}_0, \boldsymbol{e}_0, k_0, r_0^{dk}] \tag{9}$$

where $\tilde{\beta}$ is the monthly discount factor. The expectation is over the transition of future states under the chosen policy, conditional on the physician's prior beliefs about treatment-patient match values ($\boldsymbol{\theta}_0$) and initial stock of skills ($\boldsymbol{e}_0$).

Define $\beta_i = \tilde{\beta}(1 - \lambda_{i0})$ as the physician-specific discount factor that accounts for patient arrival rates, which vary significantly across physicians (Table 2). The physician receives zero payoff and does not learn when no patient arrives. Therefore I transform problem (9) for physician i to abstract away no patient arrival. The discount factor in the transformed problem is β_i, and the index t

denotes the tth patient instead of the calendar time:[15]

$$\max_{\{d_t\}_{t=0,1,\dots}} \mathbb{E}[\sum_{t=0}^{\infty} \beta_i^t u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk}) \mid \boldsymbol{\theta}_0, \boldsymbol{e}_0, k_0, r_0^{dk}] \tag{10}$$

3.3.2 Model solution: an extension to Gittins index policy

The model above features (a) a high-dimensional state space that is typical of learning models; (b) spillover effects of treatment-specific learning by doing across patient types. Once a physician's skill in one treatment improves, her future evaluation of that treatment is higher for *all* types of patients. To meet these challenges, I solve the model by extending the standard Gittins index policy in a multi-armed bandit framework.

The multi-armed bandit framework. First, the physician's problem fits naturally into a multi-armed bandit (MAB) framework. Each treatment option, $d \in \mathcal{D}$, represents an arm of the bandit machine. The arms generate random payoffs, u_t^d, that depend on their states, $(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk})$. The physician's problem (10) is to find the optimal way to operate the arms sequentially that maximizes her total expected payoff.

Fitting into the MAB framework reduces the dimensionality of the physician's problem by circumventing backward induction. Gittins (1979) proposes an index policy for the classic MAB model. He calculates for each arm an index that only depends on the arm's current state and calibrates the value of pulling it until some optimal stopping time. The Gittins index policy simply selects the arm with the highest index in each period. Gittins then shows the index policy is optimal in the standard MAB framework, which requires four assumptions: (a) exactly one arm is chosen (or *active*) in each period; (b) the unchosen arms do not generate rewards; (c) states of the unchosen arms remain frozen, generating the same average rewards in later periods; and (d) the arms are independent.

Assumptions (a) and (b) hold trivially for the physician's learning problem; (d) holds by the assumption that treatment-patient match values are independent across options; (c) is where the physician's problem deviates from the classic MAB model and falls into the realm of *restless bandits*. Restless bandits extend the classic MAB by allowing even the inactive arms to generate payoff and change states, but the stopping rule and the resulting index policy may no longer be optimal.[16]

[15] I keep the notation of t for convenience. All subsequent ts index the patient.

[16] Whittle (1988) examines generic restless bandit models in detail.

The restlessness of the physician learning model, however, solely stems from the exogenous transition of patient types (k_t). Hence it is a special case in two ways. First, unlike the generic restless bandit, the unchosen treatments do not generate any payoff. Second, the transition of patient types does not convey new information: both physician beliefs and experience $(\boldsymbol{\theta}_t^d, e_t^d)$ remain frozen for the unchosen d in period t; the restless state k only controls which of the K elements in $\boldsymbol{\theta}_t^d$ to use when the physician makes the decision.[17] Thanks to these special features, the physician learning model has an optimal solution that closely resembles the standard Gittins index policy.

The modified Gittins index policy and its optimality. Now consider an auxiliary two-armed bandit: one arm is treatment option d in state $(\boldsymbol{\theta}^d, e^d, k, \boldsymbol{r}^d)$; the other arm is the option of taking a lump-sum payoff M and retiring the arm. Denote by $\phi^d(\boldsymbol{\theta}^d, e^d, k, \boldsymbol{r}^d, M)$ the optimal expected payoff from the auxiliary bandit. Let τ^d be the time to retire, which could be $+\infty$ if d is chosen indefinitely.

Assumption 1. *Conditional on physician beliefs and skills, the expected time to retire in the two-armed bandit process is independent of the current patient's type, k:*

$$\mathbb{E}[\tau^d \mid \boldsymbol{\theta}^d, e^d, k, \boldsymbol{r}^d] = \mathbb{E}[\tau^d \mid \boldsymbol{\theta}^d, e^d, \boldsymbol{r}^d], \ \forall k \in \mathcal{K} \tag{11}$$

Note that Assumption 1 does allow the expected time to retire to depend on the physician's beliefs about treatment-patient match value, treatment-specific skills, and financial incentives. It only requires that the current patient's type has no persistent effect on the expected duration of the physician's learning process. The assumption essentially treats the realization of patient type in the initial period as a transitory shock. It is similar to assuming that the expected number of trials with d does not depend on the *order* in which patients arrive, but is even less restrictive and only focuses on the type of the *current* patient.[18]

Now I define the modified Gittins index and establish the optimality of the index policy:

Definition 1. *For arm d in state $(\boldsymbol{\theta}^d, e^d, k, \boldsymbol{r}^d)$, construct a two-armed bandit process by adding an auxiliary arm with a lump-sum retirement payment M. The Gittins index for arm d, $M^{dk}(\boldsymbol{\theta}^d, e^d, \boldsymbol{r}^d)$, is the infimum*

[17] This is different from typical restless bandits, where the evolving states of unchosen arms do change the expected payoff in future periods. See Whittle (1988) or Chapter 6 of Gittins et al. (2011).

[18] The generic restless bandit process does not always have a state variable whose effect is as transitory as the patient type in the physician's learning model, making Assumption 1 too restrictive to impose on restless bandit models in general.

of all the M values that the physician is willing to take and retire. That is

$$M^{dk}(\boldsymbol{\theta}^d, e^d, \boldsymbol{r}^d) := \inf_M \{M : \ \phi^d(\boldsymbol{\theta}^d, e^d, k, \boldsymbol{r}^d, M) = M\} \tag{12}$$

The auxiliary arm helps calibrate the Gittins index for arm d, which can be seen as a price for the sequence of payoffs from operating d *while the option of taking M^{dk} and retiring is always available* (Whittle, 1982).

Proposition 1. *Under Assumption 1, the modified Gittins index policy that always selects the treatment option with the highest $M^{dk}(\boldsymbol{\theta}^d, e^d, \boldsymbol{r}^d)$ is optimal for the physician's problem (10).*

I defer the complete proof to Appendix A. The intuition follows Whittle (1982) and is straightforward: recall that the patient type, k, is the only state variable that makes the MAB restless; the distribution of types is also exogenous, invariant over time, and known to the physician. The physician's problem from period $(t+1)$ onwards is no longer restless in expectation. Hence the modified Gittins index can evaluate current and expected future payoffs separately, and the latter resembles the standard Gittins index. Moreover, with Assumption 1, the optimal time to retire an arm is independent on the current k. Hence the optimal stopping argument for the Gittins index policy in standard MAB models goes through.

3.4 Characterization of the Gittins index

Although conceptually intuitive, the definition of the Gittins index in (12) is not helpful for calculating its value in a given state.[19] Brezzi and Lai (2002) developed a closed-form solution: they first transform the MAB problem to a Wiener process; then they apply a diffusion approximation. They show that the closed-form approximation is asymptotically optimal and performs well for short or moderate horizons as well. I adapt their results to the physician's problem and get the following approximated Gittins index:

$$\begin{aligned} M^{dk}(\boldsymbol{\theta}_t^d, e_t^d, \boldsymbol{r}_t^d) \quad = \quad & u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk}) \\ & + \beta \mathbb{E}\left[\tilde{M}(\boldsymbol{\theta}_{t+1}^d, e_{t+1}^d) + \sum_{\tau=0}^{\infty} \beta^\tau \left(\alpha \mathbb{E}(r_{t+1+\tau}^d) - c(e_{t+1+\tau}^d) \right) \middle| \boldsymbol{\theta}_t^d, e_t^d, k_t \right] \end{aligned} \tag{13}$$

The first component is the flow payoff from choosing d for the current patient with type k_t,

[19] Gittins et al. (2011) discussed various numerical methods of approximating the Gittins index with finite-horizon models, but the computational cost is still prohibitively high.

$u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, \boldsymbol{r}_t^d)$. The Gittins index has a one-to-one relationship with the flow payoff: the former must exactly compensate for any change in the latter to make the physician just indifferent between continuing with d and taking the retirement pay, M.

The second component, $\tilde{M}$, highlights the result that the physician's Bayesian learning problem becomes a standard MAB from period $(t+1)$ onward.[20] The physician expects to have a type-k patient with probability λ_k, making the *ex ante expected* success rate in any subsequent period a weighted average,

$$\bar{\theta}_{t+s}^d = \sum_{k=1}^{K} \lambda_k \theta_{t+s}^{dk}, \quad s \geq 1 \tag{14}$$

Hence the problem morphs into a MAB with one average type and the initial state $(\bar{\theta}_{t+1}^d, e_{t+1}^d, \boldsymbol{r}_t^d)$, thus getting rid of the restlessness *in expectation*. The $(t+1)$ states depend on the physician's learning result from period t: she updates her skill to $e_{t+1}^d = e_t^d + 1$; she adjusts her belief θ^{dk} upward (downward) if period t treatment is a success (failure) for $k = k_t$; the beliefs about all other types $\theta^{dk'}(k' \neq k)$ remain unchanged. The expectation in the second term of (13) is over the realization of y_t.

Now I approximate the Gittins index of the standard MAB from $(t+1)$ onward, $\tilde{M}$, following the closed-form approximation by Brezzi and Lai (2002):

$$\tilde{M}(\boldsymbol{\theta}_{t+1}^d, e_{t+1}^d) = (1-\beta)^{-1}\left[\mu_{t+1}^d + \sqrt{\nu_{t+1}^d}\,\psi\Big(\frac{\nu_{t+1}^d}{-\sigma_{\mu,d}^2 \ln \beta}\Big)\right] \tag{15}$$

where

$$\mu_{t+1}^d = \sum_{k=1}^{K} \lambda_k \mu_{t+1}^{dk}, \qquad \mu_{t+1}^{dk} = \frac{a_{t+1}^{dk}}{a_{t+1}^{dk} + b_{t+1}^{dk}} \tag{16}$$

$$\nu_{t+1}^d = \sum_{k=1}^{K} \lambda_k^2 \nu_{t+1}^{dk}, \qquad \nu_{t+1}^{dk} = \frac{a_{t+1}^{dk} b_{t+1}^{dk}}{(a_{t+1}^{dk} + b_{t+1}^{dk})^2 (a_{t+1}^{dk} + b_{t+1}^{dk} + 1)} \tag{17}$$

$$\sigma_{\mu,d}^2 = \mu_{t+1}^d (1 - \mu_{t+1}^d) \tag{18}$$

and $\psi(\cdot)$ is a closed-form, strictly increasing function with pre-calculated parameters.[21]

Finally, the last component of (13) accounts for the learning-by-doing effects and the financial incentives. The physician expects future revenues of d for each type of patient to remain at their

[20] The learning-by-doing effects and the financial incentives will be accounted for by the last term of (13).

[21] I show the complete derivation and detailed specification of ψ in Appendix B.

current level; she also takes the type-specific arrival rates into account and calculates the expected revenue for d:

$$\mathbb{E}(r^d_{t+1+\tau}) = \sum_{k=1}^{K} \lambda_k r^{dk}_{t+1+\tau} = \sum_{k=1}^{K} \lambda_k r^{dk}_t, \quad \tau \geq 0 \tag{19}$$

The physician also expects treatment cost, $c(e^d_{t+\tau})$, for clipping and coiling to keep decreasing as the she accumulates more experience. The treatment cost of observation $(d = 0)$ remains constant over time. The financial incentives and treatment costs are then summed up from $(t+1)$ onwards.

Discussion: determinants of choice probabilities. The $u^d_t(\boldsymbol{\theta}^d_t, e^d_t, k_t, \boldsymbol{r}^d)$ and μ^d_{t+1} terms of (13) show that the Gittins index is larger when the physician holds more optimistic beliefs about treatment d's (average) match value. The ν^d_{t+1} terms imply that the index is larger when the physician's beliefs are less precise, in which case the informational value of learning is higher. The β terms indicate that the index is larger when the physician discounts the future less heavily or expects more frequent patient arrivals. The $\sigma^2(\mu^d_{t+1})$ term shows that the index is larger when μ^d takes on more extreme values, and is the smaller when μ^d gets closer to 0.5. Moreover, the index is larger when the physician is not yet at the flat part of c, i.e. when the gain from learning by doing is higher.

Finally, the above effects of μ^d and ν^d on the Gittins index carry over to those of type-specific statistics, (μ^{dk}, ν^{dk}), because the former are linear combinations of the latter. For example, if the physician's belief about the match value between d and type-k is imprecise (a large ν^{dk}), then the physician has stronger incentives to experiment with procedure d. But the magnitude of such effects is governed by the type-specific arrival rates, λ_k. When type-k patients are rare, a small λ_k limits the incentive of learning on type k patients. This result is intuitive as it connects the value of learning with the probability of applying the information learned to future cases. It also echoes the third set of reduced-form evidence presented in Section 2.3.

3.5 The physician's decision rule

Given states $(\boldsymbol{\theta}^d_t, e^d_t, k_t)$, the physician makes her treatment decision in period t by solving

$$\max_d M^{dk}(\boldsymbol{\theta}^d_t, e^d_t, \boldsymbol{r}^d_t) + \xi^d_t \tag{20}$$

where ξ^d_t is an error term that is observed by the physician but unknown to the econometrician. It could stem from any idiosyncratic, treatment-specific shock such as patient preferences.[22] I assume

[22]For example, patients may be concerned about the cosmetic effects of craniotomy during surgical clipping.

that ξ_t^d follows i.i.d. Type I extreme value distribution. The approach is common in the literature and avoids degeneracy problems when forming the likelihood. It also implies simple logistic choice probabilities, thereby facilitates computation.

4 The Econometric Model and Identification

4.1 The econometric model

Heterogeneous prior beliefs. I follow Dickstein (2014) and parameterize the beta-distributed prior beliefs as

$$\begin{aligned} \theta_{i0}^{dk} \mid X_{i0}^{dk} &\sim \text{Beta}(a_{i0}^{dk}, b_{i0}^{dk}) && (21)\\ \mu(X_{i0}^{dk};\gamma_\mu) &= \frac{a_{i0}^{dk}}{a_{i0}^{dk}+b_{i0}^{dk}} = \frac{\exp\left(X_{i0}^{dk}\gamma_\mu\right)}{1+\exp\left(X_{i0}^{dk}\gamma_\mu\right)} && (22)\\ \eta(\gamma_\eta) &= a_{i0}^{dk}+b_{i0}^{dk} = \exp(\gamma_\eta) && (23) \end{aligned}$$

The prior mean, μ_{i0}^{dk}, is a logistic function of X_{i0}^{dk}: the state average adoption rate of d for type-k patients, its interaction with the dummy variable for whether physician i works primarily at a teaching hospital, its interaction with the dummy variable for whether i is a young physician, and the constant.[23] The parameterization implies the prior variance

$$\nu_{i0}^{dk} = \frac{\mu\,(1-\mu)}{1+\eta} \tag{24}$$

and the Beta distribution parameters

$$\begin{aligned} a_{i0}^{dk} &= \mu\eta && (25)\\ b_{i0}^{dk} &= (1-\mu)\,\eta && (26) \end{aligned}$$

Cost of delivering treatments, $c(\cdot)$. Following the literature on learning by doing (e.g. Argote and Epple (1990)), I assume $c(\cdot)$ is bounded, monotonically decreasing, and convex for clipping and coiling. Specifically,

$$c(e_t^d) = \begin{cases} \alpha_1^c \exp(-\alpha_2^c e_t^d), & d = 1,2 \\ \alpha_0^c, & d = 0 \end{cases} \tag{27}$$

[23] All variables in X_{i0}^{dk} are measured in period $t = 0$.

where α_1^c is the initial treatment cost before the physician has any experience; α_2^c characterizes the speed of learning, i.e. the rate at which treatment costs decreases with physician experience.[24] The cost of "delivering" watchful observation ($d = 0$) is constant at α_0^c. It is not clear whether watchful observation requires any skill from the attending physician at all, so I make the simplifying assumption that there is no learning by doing for $d = 0$.

Initial skills. I assume all physicians start with no stock of skills in coiling, i.e. $e_{i0}^2 = 0$. For clipping, I set $e_{i0}^1 = 0$ for all physicians who enter the SID panel in 2004 or later because of data limitations. Then I extrapolate the experience for the other physicians as $e_{i0}^1 = (30 - T_i) \times n_{i0}^1$, where n_{i0}^1 is physician i's caseload in 2003, and T_i is the number of years that i show up in the SID data. I assume each physician works for 30 years after completing residency and that the caseloads stay constant over time. [25]

Patient types. I group patients into 4 types along two dimensions: the type of aneurysm the patient has (ruptured or unruptured), and the patient's health condition. Ruptured and unruptured aneurysms are identified from standardized diagnosis codes (ICD-9-CM) in the SID data. Patients are considered relatively healthy if they have no major comorbidities (diseases or disorders coexistent with the primary condition, i.e. brain aneurysm) as recorded in the SID.

Type-specific patient arrival rates. I parameterize patient arrival as follows: λ_{i0} is the probability of having no patient, and is allowed to vary across physicians. λ_{ik} for types $k \in \{1, \dots, K\}$ are type-specific arrival rates. I allow them to depend on observable characteristics of physicians, namely physician tenure (young or experienced) and whether the physician works primarily at a teaching hospital.

[24] Argote and Epple (1990) show that most studies model learning in manufacturing as: $c = \alpha_1 x^{-\alpha_2}$. Compare it with the specification above, $e_t^d = \ln x$. I am essentially assuming that experience with d reduces the marginal cost of the next case *faster* than in a typical manufacturing setting, all else equal. It is likely to be the case because $c()$ captures the learning of an individual physician, who is immune to many decelerating problems with team learning, e.g. turnover and miscommunication (Reagans et al., 2005).

[25] The assumption implies that a physician retires when her record discontinues in the SID data. Thus it may overestimate the pre-2003 experience for physicians showing up in all 12 years of SID data (2003-2014).

4.2 Likelihood

Denote data on physician i in period t as ω_{it} and the model parameters as Θ. The likelihood contribution from (i, t) is

$$L_{it}(\omega_{it}; \Theta) = \begin{cases} \lambda_{i0} & \text{no patient arrives} \\ (1-\lambda_{i0})\lambda_k \prod_{d \in \mathcal{D}} \left[\dfrac{\exp(M_{it}^{dk})}{1+\sum_{d'} \exp(M_{it}^{d'k})} \left(\theta^{dk}\right)^{y_t} \left(1-\theta^{dk}\right)^{1-y_t} \right]^{d_{it}=d} & \text{otherwise} \end{cases} \tag{28}$$

where M_{it}^{dk} is specified in (13) and captures the dk-specific flow utility, the informational value of learning, and the value of cost reduction from learning. M_{it}^{dk} depends on the current states: the type of current patient, k; the physician's belief, $\boldsymbol{\theta}_{it}^{d}$, which summarizes the information on the match value of d for each type of patient she collected in the past; the physician's stock of surgical skills, e_t^d; and the expected revenues of d at the hospital where the physician receives patient t, $\boldsymbol{r}_t^d$. Note that $(\boldsymbol{\theta}_t^d, e_t^d)$ work as summary statistics for the physician's entire learning history. They affect the choice probability through the modified Gittins index, M_{it}^{dk}, which takes into account the transition of future beliefs and skills that are contingent on the current choice.

4.3 Identification

I estimate the following parameters outside the model using the SID panel data: physician-specific patient non-arrival rate, λ_{i0}; type-specific patient arrival rates that are dependent on physician characteristics, λ_{ik}; and the true match values governing treatment success rates, θ^{dk}. The richness of the SID data allows for precise estimates of these parameters.

Parameters on the treatment-type-specific prior belief determinants are identified in two steps. First, the *relative* size of the parameters are identified from the variation in initial choices across physicians with different observable characteristics, and the variation in type-specific market shares of each treatment at the time each physician started to practice. Second, *locations* of the prior means are pinned down by the differential response to patient outcomes, which is reflected in the switching patterns of physician choices in subsequent periods. Intuitively, the choice of a physician after the initial period will respond strongly to treatment failures if her prior belief is close to 1, but only moderately if her prior belief is close to 0. The reverse is true for choice responses after a success. The observable outcomes in the SID data are crucial for the second step of identification.

Parameters in the cost function for delivering treatments are identified from two sources. The scale and curvature parameters, (α_1^c, α_2^c) in $c(e_t^d) = \alpha_1^c \exp(-\alpha_2^c e_t^d)$ for clipping and coiling ($d = 1, 2$) are identified using the spillover of treatment-specific skills across patient types. Consider a physician who treated a type-k patient in period t with option $d = 1$ and receives another type-k patient in period $(t + s)$. Suppose the physician has treated *other* types of patients between periods t and $(t + s)$. The experience increases her treatment-specific surgical skill ($e_{t+s}^d > e_t^d$) but does not affect her treatment-type-specific belief ($\theta_{i,t+s}^{dk} = \theta_{it}^{dk}$). Hence the average change in the probability of choosing $d = 1$ again among similar physicians identifies the shape of the cost function. Second, the location parameter that captures the cost of observation ($d = 0$) is identified from the variation in the choice probabilities between observation relative and intervention by new physicians who have zero initial skills.

Finally, the weight on expected revenue in the physician's flow payoff is identified from the rich variation in treatment-type-specific revenues across hospitals.

5 Estimation Results

5.1 Parameter estimates

I estimate the model on the SID data from 2003 to 2014 with the maximum likelihood method. Table 8 summarizes parameters estimated directly from the data; Table 9 reports the estimates of structural parameters. For prior belief determinants, I find young and teaching-hospital physicians tend to favor intervention in general, and coiling in particular. Moreover, the prior beliefs regarding the match value of treatment d for type-k patients respond positively to the market share of d among type-k patients in the calendar quarter when a physician treated her first case. The prior beliefs of young physicians are only moderately more responsive to the initial market shares, but those of teaching hospital physicians are substantially more responsive.

The estimate on the weight of expected revenues shows that physicians have moderate financial incentives, which is consistent with findings in the literature (e.g., Johnson and Rehavi (2016)). For example, the median total charge for clipping is $20,000 higher than that for coiling. All else equal, a physician who believes the coiling success rate is 2 percentage points higher than clipping will be indifferent between the two options.

The cost to deliver clipping or coiling declines fairly rapidly under the estimates. The marginal return from learning by doing almost diminishes to zero after 15 cases. To put the numbers in perspective, a physician who has done clipping 7 times and coiling only twice will be indifferent

Patient type	Conditional arrival rates by physician characteristics (λ_k)			
	Young teaching	Young non-teaching	Experienced teaching	Experienced non-teaching
Ruptured, healthy	0.350	0.301	0.284	0.273
($k = 1$)	(0.014)	(0.010)	(0.008)	(0.006)
Ruptured, unhealthy	0.247	0.221	0.298	0.264
($k = 2$)	(0.013)	(0.009)	(0.008)	(0.006)
Unruptured, healthy	0.248	0.287	0.198	0.214
($k = 3$)	(0.013)	(0.010)	(0.007)	(0.006)
Unruptured, unhealthy	0.155	0.191	0.220	0.249
($k = 4$)	(0.011)	(0.008)	(0.007)	(0.006)

Patient type	Treatment-patient match value (θ^{dk})		
	Clip $d = 1$	Coil $d = 2$	Obs $d = 0$
Ruptured, healthy	0.271	0.400	0.377
($k = 1$)	(0.015)	(0.016)	(0.012)
Ruptured, unhealthy	0.290	0.351	0.166
($k = 2$)	(0.013)	(0.014)	(0.013)
Unruptured, healthy	0.734	0.931	0.907
($k = 3$)	(0.018)	(0.007)	(0.010)
Unruptured, unhealthy	0.591	0.843	0.849
($k = 4$)	(0.017)	(0.009)	(0.025)

Table 8: Parameters estimated directly from the SID data

NOTES: Parameters are estimated using the main sample constructed from New York SID (2003-2014). Standard errors of the mean estimator are reported in parentheses. Conditional patient arrival rate, λ_k, is the probability of getting a type-k patient in any given month, *conditional* on having any patient. Treatment-patient match value, θ^{dk}, is measured by the success rate when using treatment d on type-k patients. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Ruptured and unruptured aneurysms are identified from standardized diagnosis codes (ICD-9-CM) in the SID data. Healthy patients are those with no major comorbidities (diseases or disorders in addition to the primary condition, i.e. brain aneurysm) as recorded in the SID. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.Standard errors of the mean estimator are reported in parentheses.

Parameters	Notation	Coefficient	Std. Error
A. Determinants of prior belief (θ_{i0}^{dk})			
Constant	$\gamma_{\mu 0}$	-1.3261	0.4427
$1(d = \text{clip}) \times 1(\text{young physician})_i$	$\gamma_{\mu 1}$	0.6205	0.1316
$1(d = \text{coil}) \times 1(\text{young physician})_i$	$\gamma_{\mu 2}$	0.8853	0.2066
$1(d = \text{clip}) \times 1(\text{teaching hospital})_i$	$\gamma_{\mu 3}$	0.2267	0.0349
$1(d = \text{coil}) \times 1(\text{teaching hospital})_i$	$\gamma_{\mu 4}$	2.4203	0.2316
Initial market share of d for k	$\gamma_{\mu 5}$	0.1026	0.0134
Initial market share of d for $k \times 1(\text{young physician})_i$	$\gamma_{\mu 6}$	0.0053	0.0097
Initial market share of d for $k \times 1(\text{teaching hospital})_i$	$\gamma_{\mu 7}$	0.2354	0.0373
Prior belief imprecision	γ_{η}	5.7387	0.0013
B. Financial incentives			
Weight on total revenue (in thousand real 2014 USD)	α	0.0009	0.0001
C. Cost of delivering treatments			
Weight on surgical costs	α_1^c	0.1575	0.0665
Speed of cost reduction from learning	α_2^c	0.3761	0.1240
Cost of no intervention	α_0^c	0.0397	0.0054
$-\log(\text{Likelihood})$		12805.118	

Table 9: Structural parameter estimates

Notes: Parameters are estimated using maximum likelihood on the main sample constructed from New York SID (2003-2014). The vector of γ_μ are coefficients in the logistic function that determines the mean of physician i's prior belief about the match value of d for type-k patients. Initial market shares are the shares of type-k patients treated by d in the calendar quarter of physician i's entry. Total revenue is the total charge for the inpatient stay in thousands of real 2014 dollars. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

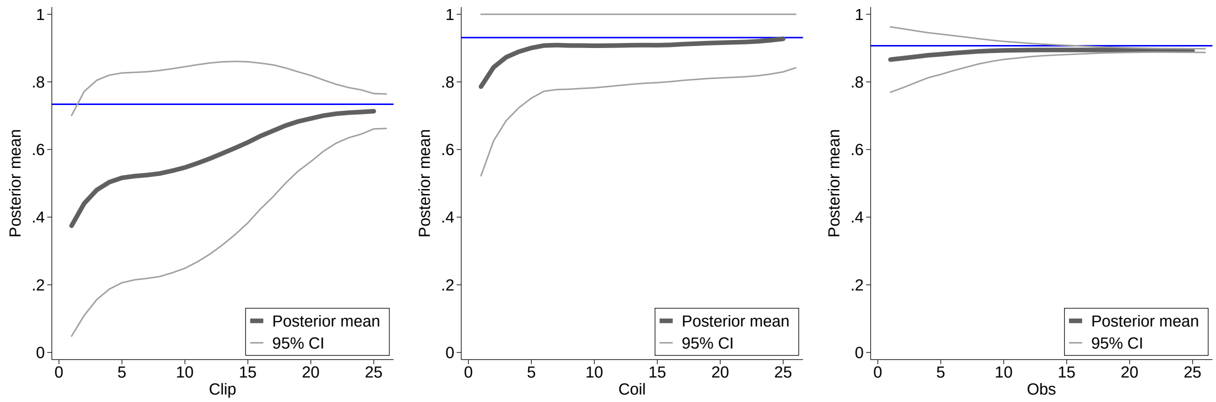


Figure 2: Example of model-implied belief evolution: converging posterior means and shrinking variances

Notes: The horizontal axes show the number of times physicians have used the treatment on the given type of patient (healthy with unruptured aneurysms). The vertical axes show beliefs about the treatment-type match value, measured by the treatment success rate. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention.The thick black line in each figure plots the posterior mean of physicians' beliefs about that treatment for the particular type of patient. The thin gray lines delineate the 95% confidence interval. The belief evolution is simulated from 1,000 samples using the estimated parameters and the observed sequences of patients. The horizontal blue line shows the "true" latent match value estimated directly from the data as a benchmark.

between the two options even if she believes the coiling success rate to be 6.5 percentage points higher.

I plot the evolution of physician posterior beliefs about the match values of the three options for healthy patients with unruptured aneurysms as an example. Figure 2 traces the simulate posterior beliefs against the number of times physicians have used the treatment on that type of patient. The horizontal blue line shows the latent match value estimated directly from the data as a benchmark. The thick black line in each figure plots the posterior mean of physicians' beliefs about that treatment for the particular type of patient, which converges rapidly to the "truth." The thin gray lines delineate the 95% confidence interval, which shrinks over time.

5.2 Model fit: choice probabilities and transition dynamics

I first compare the observed and model-predicted choice probabilities as a first evaluation of model fit. Table 10 compares the observed and predicted choice probabilities, and Figures 3 and 4 illustrate the comparison. The model captures the overall pattern, with slightly higher choices probabilities for the two interventions. It also does a good job in fitting the patient type-specific choices probabilities for 2 of the 4 types. For patients with unruptured aneurysms, the model overpredicts the probability of no intervention on healthy patients and the probability of clipping on unhealthy ones. Note that the direct comparison of choices probabilities is a stringent assessment of the model fit. The strong learning-related path dependence tends to exacerbate any difference in observed and predicted choices in the earlier periods.

		Choice probability		
		Clipping	Coiling	Observation
Overall	Data	0.303	0.409	0.288
	Model	[0.342]	[0.339]	[0.320]
By patient type				
Ruptured, healthy	Data	0.251	0.275	0.474
	Model	[0.237]	[0.282]	[0.481]
Ruptured, unhealthy	Data	0.401	0.351	0.248
	Model	[0.487]	[0.313]	[0.200]
Unruptured, healthy	Data	0.236	0.466	0.299
	Model	[0.131]	[0.358]	[0.511]
Unruptured, unhealthy	Data	0.325	0.594	0.082
	Model	[0.519]	[0.425]	[0.056]

Table 10: Model fit: choice probabilities by patient type
NOTES: The *Data* rows report observed choice probabilities in the SID data. The *Model* rows report model-predicted choice probabilities (in brackets) from 1,000 simulations using the estimated parameters and the observed sequences of patients. Ruptured and unruptured aneurysms are identified from standardized diagnosis codes (ICD-9-CM) in the SID data. Healthy patients are those with no major comorbidities (diseases or disorders in addition to the primary condition, i.e. brain aneurysm) as recorded in the SID.

I then compare the observed and model-predicted choice *transition* probabilities conditional on previous period choices, outcomes, and patient types. Table 11 summarizes the results; Figure 5 visualizes the comparison.

The model predictions closely resemble what is observed in the data. I first examine the overall probability of choosing the same treatment in period $(t+1)$ as that in period d_t, conditional on the previous period choice (d_t), outcome (y_t), and whether patient types in the two periods are the same. The conditional probabilities capture the *reluctance* of physicians to experiment further with a treatment option.

Both the data and the model predictions show that physicians are more likely to keep using a treatment when the previous case treated with it was a success, especially when the patients in the two adjacent periods are of the same type. This is consistent with the relevance of Bayesian learning only for the same type, because beliefs about treatment match values are type-specific.

I further break down the whole sample by physician skill level. Table 12 shows the transition probabilities when the physician's experience e_t^d is high (greater than the 75th percentile), medium

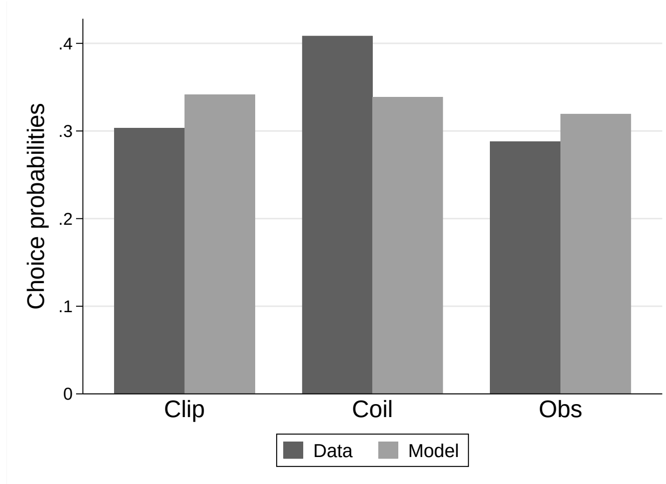


Figure 3: Model fit: overall choice probabilities
Notes: The *Data* bars plot the average choice probability of a treatment observed in the data. The *Model* bars plot the predicted probability simulated from 1,000 samples using the estimated parameters and the observed sequences of patients. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention.

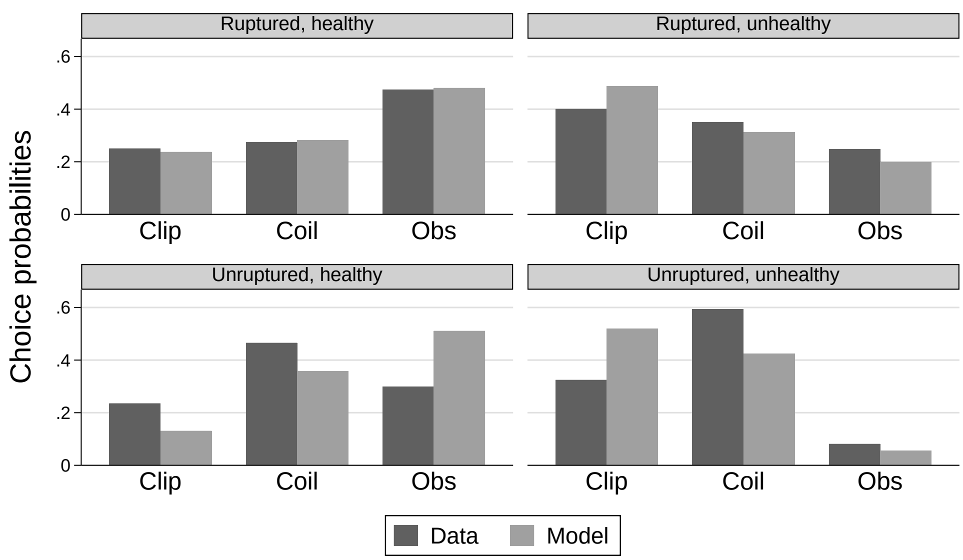


Figure 4: Model fit: choice probabilities by patient type
Notes: The *Data* bars plot the average choice probability of a treatment for each of the four patient types observed in the data. The *Model* bars plot the predicted probability simulated from 1,000 samples using the estimated parameters and the observed sequences of patients. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Ruptured and unruptured aneurysms are identified from standardized diagnosis codes (ICD-9-CM) in the SID data. Healthy patients are those with no major comorbidities (diseases or disorders in addition to the primary condition, i.e. brain aneurysm) as recorded in the SID.

	$\Pr(d_{t+1} = d_t \mid d_t, y_t, k_t, k_{t+1})$		
	Clip	Coil	Obs
	(1)	(2)	(3)
Same type, previous case was a success			
$\Pr(d_{t+1} = d_t \mid d_t, y_t = 1, k_t = k_{t+1})$			
Data	0.574	0.761	0.703
Model	[0.633]	[0.703]	[0.684]
Same type, previous case was a failure			
$\Pr(d_{t+1} = d_t \mid d_t, y_t = 0, k_t = k_{t+1})$			
Data	0.553	0.541	0.607
Model	[0.597]	[0.622]	[0.564]
Different type, previous case was a success			
$\Pr(d_{t+1} = d_t \mid d_t, y_t = 1, k_t \neq k_{t+1})$			
Data	0.464	0.591	0.350
Model	[0.450]	[0.671]	[0.312]
Different type, previous case was a failure			
$\Pr(d_{t+1} = d_t \mid d_t, y_t = 0, k_t \neq k_{t+1})$			
Data	0.463	0.524	0.468
Model	[0.445]	[0.652]	[0.355]

Table 11: Model fit: choice transition probabilities conditional on preceding choice, patient type, and outcome

NOTES: The table reports the probability of choosing the same treatment in period $(t+1)$ as that in period d_t, conditional on the previous period choice (d_t), outcome (y_t), and whether patient types in the two periods are the same. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention.*Data* rows show observed probabilities; *Model* rows show model-predicted probabilities (in brackets) from 1,000 simulations using the estimated parameters and the observed sequences of patients.

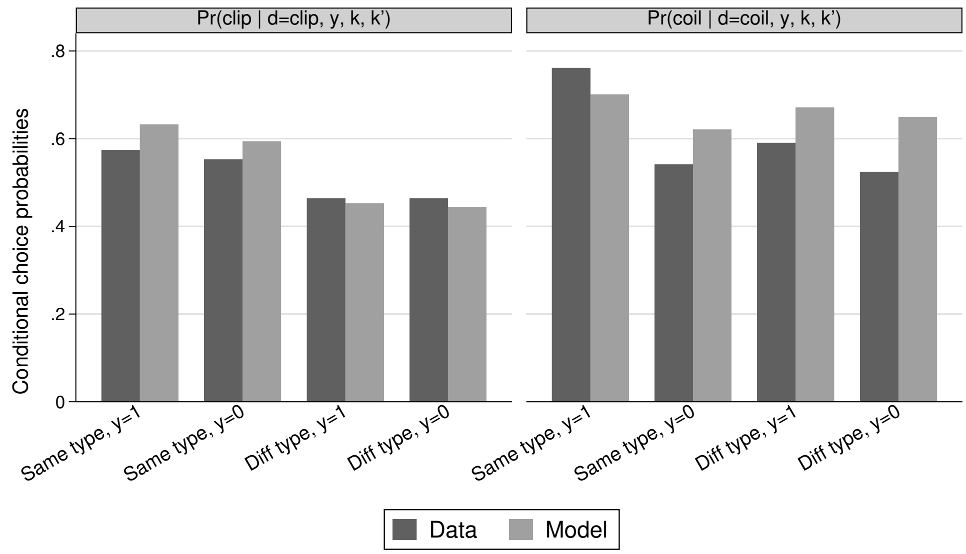


Figure 5: Model fit: choice transition probabilities conditional on preceding choice, patient type, and outcome

Notes: The *Data* bars in each figure plot the probability of choosing *the same* treatment in a period as the previous period, conditional on whether patients in the two periods are of the same type, and whether the previous treatment is a success ($y = 1$). A treatment is a success if the patient can be discharged home and does not need assisted care.The left panel shows this probability for clipping, and the right for coiling. The *Model* bars plot the predicted probability simulated from 1,000 samples using the estimated parameters and the observed sequences of patients.

(below the 75th percentile and above the 25th percentile), or low (below the 25th percentile). Figure 6 illustrates the comparison. The predicted choice transition probabilities are able to capture the stronger tendency to keep choosing the same treatment when the physicians already have high skills in it. The choice patterns highlight the diminishing learning by doing incentives as the surgical skills accumulate.

5.3 Disentangling Bayesian learning from learning by doing

With the model estimates, I use simulation to disentangle the impacts of Bayesian learning and learning by doing. Because of interactions between the two kinds of learning, their implications on treatment choices and outcomes are often hard to separate in reduced-form analyses.

Table 13 compares the observed choice probabilities and patient outcomes with those when one or both kinds of learning are shut down. The first row (*Data*) in the top panel shows the overall choice probabilities in the data. The second row (*No Bayesian learning*) reports the choice probabilities from 1,000 simulations based on the model estimates, assuming physicians only accumulate skills via learning by doing but no longer update beliefs via Bayesian learning. The resulting

	$\Pr(d_{t+1} = d_t \mid d_t, y_t, k_t, k_{t+1})$								
	High experience			Medium experience			Low experience		
	Clip	Coil	Obs	Clip	Coil	Obs	Clip	Coil	Obs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Same type, previous case was a success: $\Pr(d_{t+1} = d_t \mid d_t, y_t = 1, k_t = k_{t+1})$									
Data	0.635	0.792	0.662	0.566	0.740	0.758	0.453	0.627	0.625
Model	[0.683]	[0.795]	[0.753]	[0.604]	[0.653]	[0.600]	[0.582]	[0.562]	[0.606]
Same type, previous case was a failure: $\Pr(d_{t+1} = d_t \mid d_t, y_t = 0, k_t = k_{t+1})$									
Data	0.628	0.681	0.563	0.575	0.515	0.652	0.375	0.382	0.594
Model	[0.662]	[0.769]	[0.587]	[0.574]	[0.589]	[0.549]	[0.537]	[0.450]	[0.560]
Different type, previous case was a success: $\Pr(d_{t+1} = d_t \mid d_t, y_t = 1, k_t \neq k_{t+1})$									
Data	0.486	0.668	0.411	0.484	0.563	0.331	0.369	0.311	0.287
Model	[0.422]	[0.867]	[0.213]	[0.484]	[0.673]	[0.303]	[0.520]	[0.432]	[0.340]
Different type, previous case was a failure: $\Pr(d_{t+1} = d_t \mid d_t, y_t = 0, k_t \neq k_{t+1})$									
Data	0.516	0.635	0.484	0.468	0.520	0.504	0.363	0.221	0.339
Model	[0.442]	[0.862]	[0.467]	[0.439]	[0.681]	[0.350]	[0.489]	[0.416]	[0.361]

Table 12: Model fit: choice transition probabilities conditional on preceding choice, patient type, and outcome (by physician skill level)

NOTES: The table reports the probability of choosing the same treatment in period $(t+1)$ as that in period d_t, conditional on the previous period choice (d_t), outcome (y_t), and whether patient types in the two periods are the same. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. *Data* rows show observed probabilities; *Model* rows show model-predicted probabilities (in brackets) from 1,000 simulations using the estimated parameters and the observed sequences of patients. The first three columns report the transition probability for the whole sample. The remaining 9 columns show the transition probabilities when the physician's experience e_t^d is high (greater than the 75th percentile), medium (below the 75th percentile and above the 25th percentile), or low (below the 25th percentile).

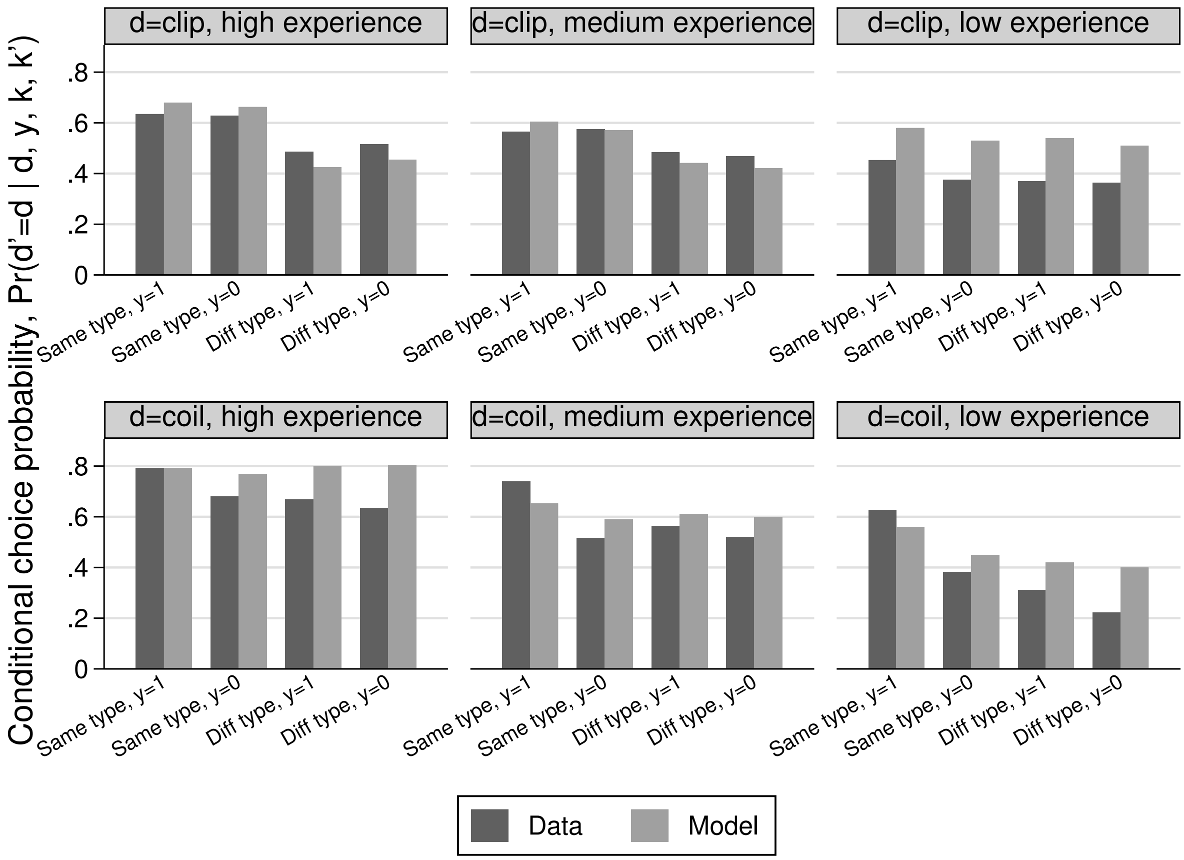


Figure 6: Model fit: choice transition probabilities conditional on preceding choice, patient type, and outcome (by physician skill level)

Notes: The *Data* bars in each figure plot the probability of choosing *the same* treatment in a period as the previous period, conditional on whether patients in the two periods are of the same type, and whether the previous treatment was a success ($y = 1$). A treatment is a success if the patient can be discharged home and does not need assisted care.The top 3 panels report the probability for clipping, and the bottom 3 for coiling. The *Model* bars plot the predicted probability simulated from 1,000 samples using the estimated parameters and the observed sequences of patients. The 3 columns use the subsamples of physicians whose experience e_t^d is high (greater than the 75th percentile), medium (below the 75th percentile and above the 25th percentile), or low (below the 25th percentile), respectively.

choice probability of clipping is 48.3%, a sizable increase from the observed 30.3%; that of coiling decreases to 34.4% from the observed 40.9%. The corresponding patient outcomes also worsens: the overall treatment success rate is 52.6%, a 1.8 percentage points reduction from the observed 54.4%.

The third row (*No learning*) further assumes that physicians no long accumulate skills via learning by doing as well, thereby blocking both kinds of learning. The simulated choice probability of clipping increases even more to 53.4%, while that of coiling more than halves from the previously level to a mere 16.3%. Treatment success rate also decreases to 51%, 3.4 percentage points lower than that observed in the data.

The comparison disentangles the effects of Bayesian learning and learning by doing. The effect of Bayesian learning dominates that of learning by doing on the utilization of clipping, explaining 78% of the changes when learning is shut down; but it is dominated by learning by doing on the adoption of coiling, explaining only 26% of the changes. Their effects on patient outcomes are comparable in magnitudes, with Bayesian learning's share being 54% and learning by doing's being 46%.

The relative magnitudes of Bayesian learning and learning by doing incentives also vary substantially across different types of physicians. The remainder of Table 13 breaks down the comparison by physician subgroups. For example, experienced physicians at non-teaching hospitals demonstrate a stronger tendency to choose clipping instead of coiling when their learning incentives are removed. This is consistent with their higher initial stock of surgical skills in clipping, as well as less optimistic prior beliefs in coiling.

6 Counterfactual Analyses

I conduct two sets of counterfactual analyses in this section. First, I quantify the magnitude of physician learning. To that end, I (a) explore how choices and outcomes would change if physicians were myopic and ignore the future returns to learning; and (b) measure how the variation in physician choices would change if one or both channels of learning are shut down. Second, I evaluate the impacts of two payment reforms: uniform payments regardless of the choice of intervention; and outcome-based payments that reward good patient outcomes and penalize bad ones.

	Choice probability			Success
	Pr(clip)	Pr(coil)	Pr(obs)	rates
Overall				
Data	0.303	0.409	0.288	0.544
No Bayesian learning	0.483	0.344	0.173	0.526
No learning	0.534	0.163	0.304	0.510
By physician subgroups				
Experienced physicians at non-teaching hospitals				
Data	0.307	0.341	0.353	0.528
No Bayesian learning	0.579	0.138	0.283	0.509
No learning	0.640	0.048	0.312	0.496
Experienced physicians at teaching hospitals				
Data	0.293	0.431	0.277	0.557
No Bayesian learning	0.427	0.475	0.098	0.514
No learning	0.711	0.092	0.198	0.465
Young physicians at non-teaching hospitals				
Data	0.334	0.460	0.206	0.559
No Bayesian learning	0.408	0.503	0.090	0.562
No learning	0.302	0.372	0.326	0.565
Young physicians at teaching hospitals				
Data	0.242	0.498	0.260	0.541
No Bayesian learning	0.425	0.450	0.125	0.531
No learning	0.236	0.282	0.483	0.541

Table 13: Simulation analyses: disentangling the effects of two kinds of learning

NOTES: *No Bayesian learning* and *No learning* rows report the simulated choice probabilities and patient outcomes assuming there is only learning by doing but no Bayesian learning, and no learning at all, respectively. The results are based on 1,000 simulations using the estimated parameters and the observed sequences of patients. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

6.1 Physician myopia

Physicians could be myopic in two ways: they could make treatment choices solely based on the expected outcome of the current patient; or they could make treatment choices based on their flow payoff, taking into account the current patient's outcome, the expected revenue, and the cost of delivering the chosen treatment.

Table 14 compares the choice probabilities and patient outcomes, assuming physicians maximize the expected outcome of their current patients. With this form of myopia, physicians would choose clipping less often than observed in the data. The contrast suggests that the observed choices of clipping may be partially driven by lower costs of delivering treatment due to high stocks of clipping skills, or by the higher expected revenue of clipping. The change in the choice probabilities of coiling varies across different groups of physicians: young physicians at teaching hospitals would choose coiling more often, reflecting the above-noted higher prior beliefs about the value of coiling; other types of physicians choose coiling less often. But all groups of physicians would choose no intervention more often had they been solely maximizing patient outcomes, which suggests over-intervention in the data.

The over-intervention is often at the expense of inferior patient outcomes. Panel B of Table 14 shows that all types of patients would have moderately higher chances of good outcomes if physicians were myopic and maximizing patient outcomes. Moreover, patients of the two unhealthy types are most likely to receive a different treatment and would see more improvement in outcomes than with physician myopia. This suggests that forward-looking physicians tend to deviate more from the myopic best treatment on unhealthy types, thereby hurting the outcomes of these patients.

The deviation from myopic best choices does not imply inferior outcomes for all patients, though. The right panel of Figure 7 compares the observed outcomes (dark bars) with those under outcome maximization (light bars). The outcomes would be better under physician myopia *on average*, but the potential changes vary. For patients observed to receive clipping, the outcomes would improve significantly if physicians were myopic. For patients observed to receive coiling, however, the outcomes would be slightly worse under myopia. This shows that experimentation can help physicians learn about the match value of some treatments, thereby improving patient outcomes.

That physicians only maximize the expected outcome of current patients is in stark contrast with strategic forward-looking physicians maximizing their total discounted payoffs. Hence I also

	Choice probability			Success	Fraction with
	Pr(clip)	Pr(coil)	Pr(obs)	rates	deviation
			Overall		
Data	0.303	0.409	0.288	0.544	
Model	0.222	0.399	0.379	0.573	
			A. By physician subgroups		
Experienced physicians at non-teaching hospitals					
Data	0.307	0.341	0.353	0.528	
Model	0.270	0.263	0.466	0.579	
Experienced physicians at teaching hospitals					
Data	0.293	0.431	0.277	0.557	
Model	0.234	0.341	0.425	0.536	
Young physicians at non-teaching hospitals					
Data	0.334	0.460	0.206	0.559	
Model	0.127	0.659	0.214	0.605	
Young physicians at teaching hospitals					
Data	0.242	0.498	0.260	0.541	
Model	0.228	0.448	0.323	0.561	
			B. By patient types k		
Ruptured, healthy					
Data	0.251	0.275	0.474	0.357	
Model	0.116	0.185	0.699	0.369	0.301
Ruptured, unhealthy					
Data	0.401	0.351	0.248	0.281	
Model	0.633	0.276	0.092	0.295	0.908
Unruptured, healthy					
Data	0.236	0.466	0.399	0.877	
Model	0.009	0.318	0.673	0.913	0.327
Unruptured, unhealthy					
Data	0.325	0.594	0.082	0.761	
Model	0.087	0.909	0.004	0.821	0.996

Table 14: Counterfactual experiment: physicians maximizing current patient outcomes

Notes: The *Model* rows report the simulated choice probabilities and patient outcomes, assuming physicians maximize current-period expected patient outcomes (plus a treatment-specific, Type-I extreme value error). The results are based on 1,000 simulations using the estimated parameters and the observed sequences of patients. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention.*Fraction with deviation* is the fraction of patients with the given type who would receive a different treatment had the physician been maximizing patient outcome. Panel A breaks the whole sample down by physician subgroups; Panel B by patient types. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used. Ruptured and unruptured aneurysms are identified from standardized diagnosis codes (ICD-9-CM) in the SID data. Healthy patients are those with no major comorbidities (diseases or disorders in addition to the primary condition, i.e. brain aneurysm) as recorded in the SID.

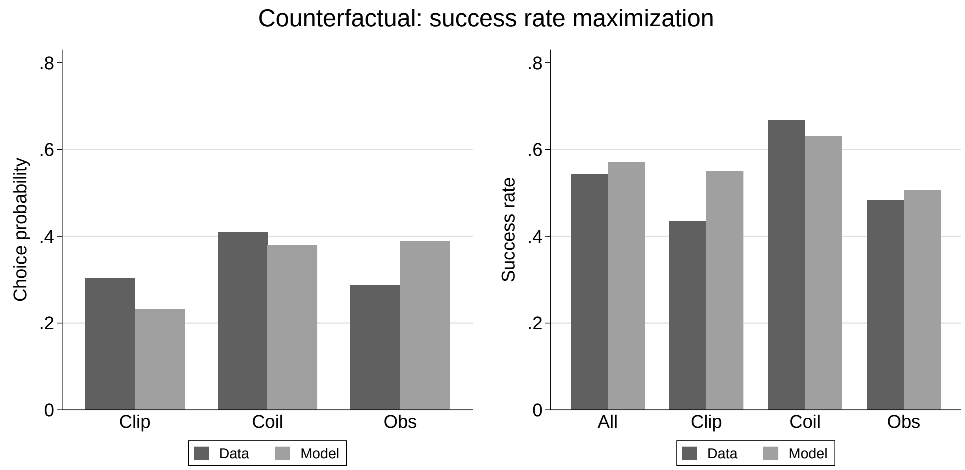


Figure 7: Counterfactual experiment: choice probabilities when physicians maximize current patient outcomes

NOTES: The *Data* bars plot the observed choice probabilities (left panel) and treatment success rates (right panel). The *Model* bars plot the simulated choice probabilities and patient outcomes, assuming physicians maximize current-period expected patient outcomes (plus a treatment-specific, Type-I extreme value error). *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. A treatment is a success if the patient can be discharged home and does not need assisted care.

examine the choices and outcomes if physicians only maximize their flow payoff but ignore the continuation values. Table 15 compares the observed and simulated choices and outcomes. On average, the choice probabilities for clipping stay relatively unchanged, but those for coiling decrease substantially if physicians were only maximizing flow payoffs. Additionally, there is a lower probability that patient outcomes would be good. The differences imply that it is the future value of learning that encourages experimentation with coiling, which helps improve patient outcomes.

Panel A of Table 15 breaks the whole sample down by physician groups. All groups of physicians would reduce the adoption of coiling under myopia. Moreover, experienced physician choices will mainly transition towards clipping, in which they had accumulated more skills; young physician choices will mainly transition towards no intervention, given their low skills in both clipping and coiling.

Panel B of Table 15 looks at the differential response by patient type. Physicians still tend to deviate from the treatment that maximizes the flow utility when seeing patients of unhealthy types. The difference between observed and predicted treatment success rates also shows that these types of patients actually benefit from the experimentation of forward-looking physicians: their success rates would be 4-9 percentage points lower had the physician ignored the future values of learning.

Figure 8 illustrates the changes in choice probabilities and patient outcomes if physicians only

	Choice probability			Success	Fraction with
	Pr(clip)	Pr(coil)	Pr(obs)	rates	deviation
			Overall		
Data	0.303	0.409	0.288	0.544	
Model	0.308	0.071	0.548	0.516	
			A. By physician subgroups		
Experienced physicians at non-teaching hospitals					
Data	0.307	0.341	0.353	0.528	
Model	0.502	0.013	0.485	0.508	
Experienced physicians at teaching hospitals					
Data	0.293	0.431	0.277	0.557	
Model	0.515	0.010	0.475	0.468	
Young physicians at non-teaching hospitals					
Data	0.334	0.460	0.206	0.559	
Model	0.153	0.209	0.639	0.568	
Young physicians at teaching hospitals					
Data	0.242	0.498	0.260	0.541	
Model	0.116	0.121	0.762	0.536	
			B. By patient types k		
Ruptured, healthy					
Data	0.251	0.275	0.474	0.357	
Model	0.235	0.052	0.713	0.353	0.287
Ruptured, unhealthy					
Data	0.401	0.351	0.248	0.281	
Model	0.519	0.053	0.428	0.240	0.572
Unruptured, healthy					
Data	0.236	0.466	0.399	0.877	
Model	0.117	0.090	0.793	0.889	0.207
Unruptured, unhealthy					
Data	0.325	0.594	0.082	0.761	
Model	0.675	0.098	0.227	0.674	0.773

Table 15: Counterfactual experiment: physicians maximizing flow payoff

Notes: The *Model* rows report the simulated choice probabilities and patient outcomes assuming physicians maximize flow payoff (plus a treatment-specific, Type-I extreme value error). The results are based on 1,000 simulations using the estimated parameters and the observed sequences of patients. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention.*Fraction with deviation* is the fraction of patients with the given type who would receive a different treatment had the physician been maximizing her flow payoff. Panel A breaks the whole sample down by physician subgroups; Panel B by patient types. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used. Ruptured and unruptured aneurysms are identified from standardized diagnosis codes (ICD-9-CM) in the SID data. Healthy patients are those with no major comorbidities (diseases or disorders in addition to the primary condition, i.e. brain aneurysm) as recorded in the SID.

maximize their own flow payoffs. Patients would get inferior outcomes on average, which is the artifact of the physicians taking into account the cost of delivering care and expected revenue in addition to patient outcomes. Patients who are observed to receive coiling see the largest decline in outcomes with physician myopia. The decline highlights the value of learning by forward-looking physicians.

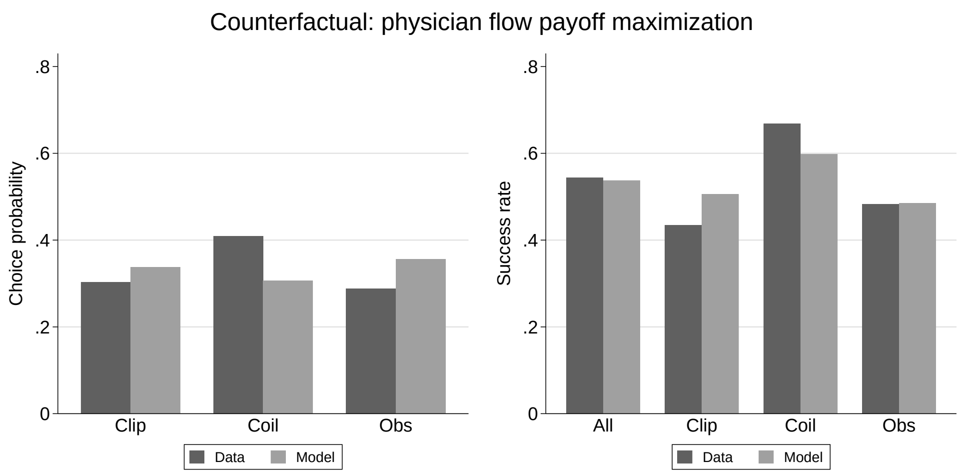


Figure 8: Counterfactual experiment: choice probabilities when physicians maximize flow payoff
Notes: The *Data* bars plot the observed choice probabilities (left panel) and treatment success rates (right panel). The *Model* bars plot the simulated choice probabilities and patient outcomes, assuming physicians maximize flow payoff (plus a treatment-specific, Type-I extreme value error). *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. A treatment is a success if the patient can be discharged home and does not need assisted care.

Finally, Figure 9 summarizes the heterogeneous effects on outcomes by patient type. Not surprisingly, all patients would get better outcomes if physicians were maximizing the current patient outcomes. But when physicians maximize their own flow payoffs, patients of unhealthy types would have lower treatment success rates because of the reduced experimentation. The outcomes of healthy patients would not change much.

6.2 Learning and variations in the choice of care

I also gauge the magnitude of learning by quantifying its contribution to the variation in physicians' choice of care. The variation of particular interest is that in the physicians' utilization rate of coiling, the new procedure, where learning tend to be more important. To this end, I shut down one or both learning channels and compare the resulting overall, within-, and between-physician variation.

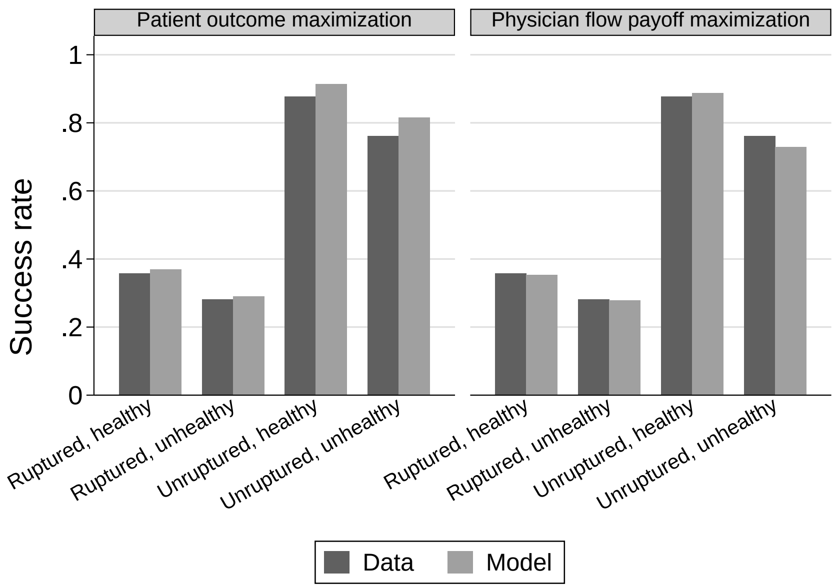


Figure 9: Counterfactual experiment: patient outcomes under physician myopia

NOTES: The *Data* bars plot the observed treatment success rates. The *Model* bars plot the simulated success rates under physician myopia. A treatment is a success if the patient can be discharged home and does not need assisted care.The right panel assumes physicians maximize current patient outcomes, and the left panel assumes physicians maximize their own flow payoffs. Ruptured and unruptured aneurysms are identified from standardized diagnosis codes (ICD-9-CM) in the SID data. Healthy patients are those with no major comorbidities (diseases or disorders in addition to the primary condition, i.e. brain aneurysm) as recorded in the SID.

Panel A of Table 16 presents the baseline variation in physicians choice probability of coiling. The first row shows a substantial overall variation of 0.292 among all cases in the sample; the between-physician and within-physician variations are 0.199 and 0.116, respectively. The four subsequent rows break the sample down by patient type. Physicians vary more in their choice of care for unruptured cases, conditional on patient health, and for relatively healthy patients, conditional on the type of aneurysm.

Panel B shows the counterfactual variation when only learning by doing is present, while Bayesian learning is shut down: Physicians still start from heterogeneous priors, but no longer update their beliefs after observing patient outcomes. Although the resulting overall and between-physician variation shrink only slightly compared with the baseline, the within-physician increases by at least 20%. That is, Bayesian learning helps to reduce within-physician variation as physician beliefs converge.

Conversely, Panel C keeps only Bayesian learning and shuts down learning by doing: Physicians' cost of delivering surgery no longer decreases as they build volume. Now it is the between-physician that changes significantly, with a 23% reduction in the whole sample. The within-physician variation, on the contrary, increases by varying amounts depending on the patient type. The overall variation is dominated by the former effect, though, with reductions between 16-25% across different patient types. Hence learning by doing mainly adds to the within-physician variation. The intuition is also strong here: learning by doing solidifies the randomness in a physician's past outcomes, thereby creating divergence even among physicians were initially similar.

Finally, Panel D shuts down both types of learning. The overall variation shrinks by 20%, which is the fraction that is attributable to learning. Previous studies (such as Finkelstein et al. (2016)) have found that supply-side factors explain about 60% of total variation in the choice of care. Taking that figure into account, learning explains about one-third of supply-side variation, which is a considerable share.

6.3 Payment reforms

I continue to use the estimated model to evaluate the potential impacts of two payment reforms. I first examine a payment reform where all payments are independent of the treatment choices, except when no intervention is chosen. Then I examine a reform that links payments with patient outcomes, using the Value Modifier (VM) program by the Centers for Medicare of Medicaid Services (CMS) as an example.

	Overall variation	Change*	Between-physician	Change*	Within-physician	Change*
A. Baseline (Learning by doing and Bayesian learning both present)						
All	0.292		0.199		0.116	
Ruptured, healthy	0.278		0.215		0.071	
Ruptured, unhealthy	0.249		0.187		0.094	
Unruptured, healthy	0.335		0.248		0.057	
Unruptured, unhealthy	0.287		0.191		0.104	
B. Only learning by doing present; Bayesian learning shut down						
All	0.282	-3.4%	0.193	-3.3%	0.140	21.1%
Ruptured, healthy	0.273	-1.8%	0.213	-1.1%	0.105	47.8%
Ruptured, unhealthy	0.250	0.2%	0.183	-2.5%	0.126	34.7%
Unruptured, healthy	0.330	-1.3%	0.253	2.2%	0.117	107.2%
Unruptured, unhealthy	0.268	-6.7%	0.173	-9.0%	0.125	20.2%
C. Only Bayesian learning present; learning by doing shut down						
All	0.230	-21.0%	0.154	-23.0%	0.123	5.6%
Ruptured, healthy	0.208	-25.1%	0.157	-26.8%	0.097	36.5%
Ruptured, unhealthy	0.200	-19.9%	0.164	-12.3%	0.104	11.5%
Unruptured, healthy	0.272	-18.6%	0.182	-26.6%	0.107	88.4%
Unruptured, unhealthy	0.242	-15.8%	0.183	-3.8%	0.109	4.7%
D. Neither kinds of learning present, both shut down						
All	0.234	-19.8%	15.9%	-20.4%	0.116	0.0%
Ruptured, healthy	0.212	-23.8%	0.163	-24.4%	0.087	22.5%
Ruptured, unhealthy	0.202	-19.0%	0.169	-9.9%	0.097	3.9%
Unruptured, healthy	0.278	-16.9%	0.184	-25.5%	0.105	85.3%
Unruptured, unhealthy	0.244	-14.8%	0.187	-1.7%	0.101	-2.9%

Table 16: Counterfactual experiment: effects of learning on choice variation

Notes: * *Change* reports the change in the counterfactual variation from the baseline, as a percentage of the baseline variation. Variation in the choice of care is measured by the standard deviation in the fraction of cases treated by coiling. *All* refers to the whole sample; *Ruptured, healthy* and similar rows use only the subsample of cases of that type.

6.3.1 Uniform payments across treatment

As noted above, the payment differentials between clipping and coiling may induce physicians to deviate from the treatment that maximizes patient outcomes. As such, I explore the change in physician choice probabilities and patient outcomes if payments are independent of the choice of intervention. In this counterfactual experiment, I set the total charges for clipping and coiling to be the same within each hospital, conditional on patient type. The charges for no intervention are left unchanged.

Table 17 shows that when clipping and coiling generate the same revenue, physicians tend to choose more intervention in general. The resulting patient outcomes are similar, but physician revenues per case increase substantially from 114 to 176 thousand dollars. The physician responses vary by tenure and the type of hospital, though. Experienced physicians at non-teaching hospitals tend to do even more clipping and less coiling when the revenues are independent of choices. This implies that the existing fee differentials between clipping and coiling for these physicians are counteracting the differences in physician beliefs and skills that favor clipping. For other groups of physicians, however, the payment reform induces more take-up of coiling, suggesting that the existing fee differentials are discouraging the choice of coiling.

Figures 10 and 11 illustrate the comparisons of choice probabilities and patient outcomes, respectively. Figure 10 highlights the heterogeneous response across physician groups to the payment reform: all physicians except those who are experienced and work at non-teaching hospitals would choose coiling more often when the fee differentials are eliminated. Figure 11 summarizes the change in patient outcomes, breaking the whole sample down by both physician groups and observed treatments in the data. Overall, patients treated by experienced physicians would have slightly lower success rates under the payment reform, and those treated by young physicians would have superior outcomes. Between patients who are observed to receive different treatments, those treated with clipping in the data would generally have better outcomes had the payments been equal. On the contrary, those treated with coiling in the data would have moderately worse outcomes under uniform payments.

6.3.2 Outcome-contingent payments

There have been ongoing proposals to link payments for medical services with patient outcomes. The CMS is currently rolling out the Value Modifier (VM) program for Medicare bene-

	Choice probability			Success	Revenue/case
	Pr(clip)	Pr(coil)	Pr(obs)	rates	($1,000)
			Overall		
Data	0.303	0.409	0.288	0.544	114.25
Model	0.366	0.472	0.162	0.545	176.06
			By physician subgroups		
Experienced physicians at non-teaching hospitals					
Data	0.307	0.341	0.353	0.528	87.96
Model	0.515	0.211	0.274	0.521	138.00
Experienced physicians at teaching hospitals					
Data	0.293	0.431	0.277	0.557	119.00
Model	0.324	0.591	0.085	0.530	203.45
Young physicians at non-teaching hospitals					
Data	0.334	0.460	0.206	0.559	131.16
Model	0.207	0.713	0.080	0.594	184.56
Young physicians at teaching hospitals					
Data	0.242	0.498	0.260	0.541	165.58
Model	0.261	0.631	0.108	0.557	238.07

Table 17: Counterfactual experiment: uniform payments for clipping and coiling

Notes: The *Model* rows report the simulated choice probabilities and patient outcomes, assuming uniform payments between clipping and coiling, conditional on patient types and hospitals. The results are based on 1,000 simulations using the estimated parameters and the observed sequences of patients. Revenue per case is in thousands of real 2014 dollars. A treatment is a success if the patient can be discharged home and does not need assisted care. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

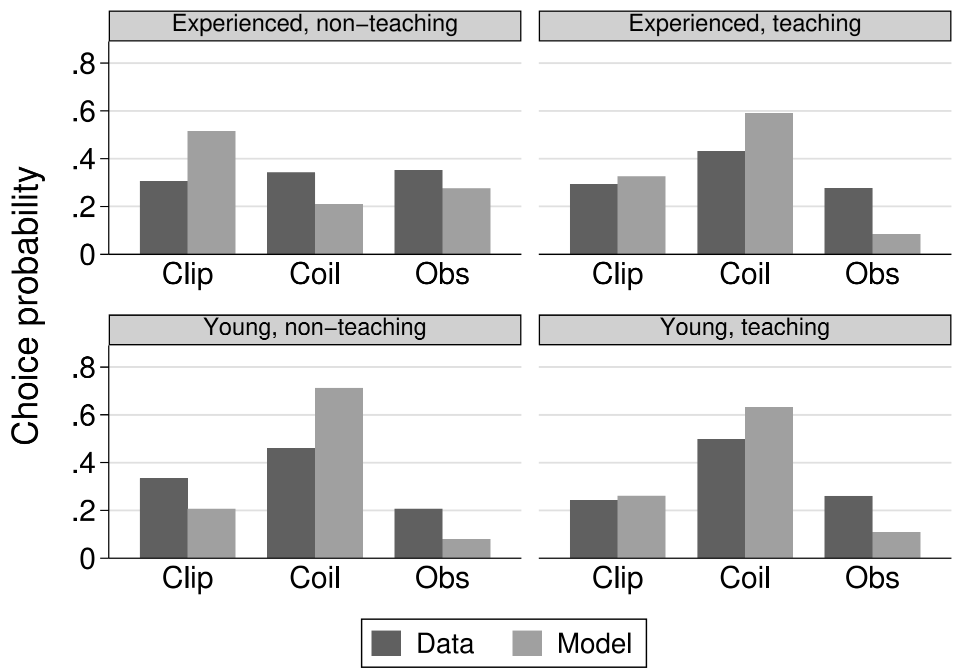


Figure 10: Counterfactual experiment: choice probabilities under uniform payments for clipping and coiling

Notes: The *Data* bars plot the observed choice probabilities by physician groups. The *Model* bars plot the simulated probabilities, assuming uniform payments between clipping and coiling, conditional on patient types and hospitals. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

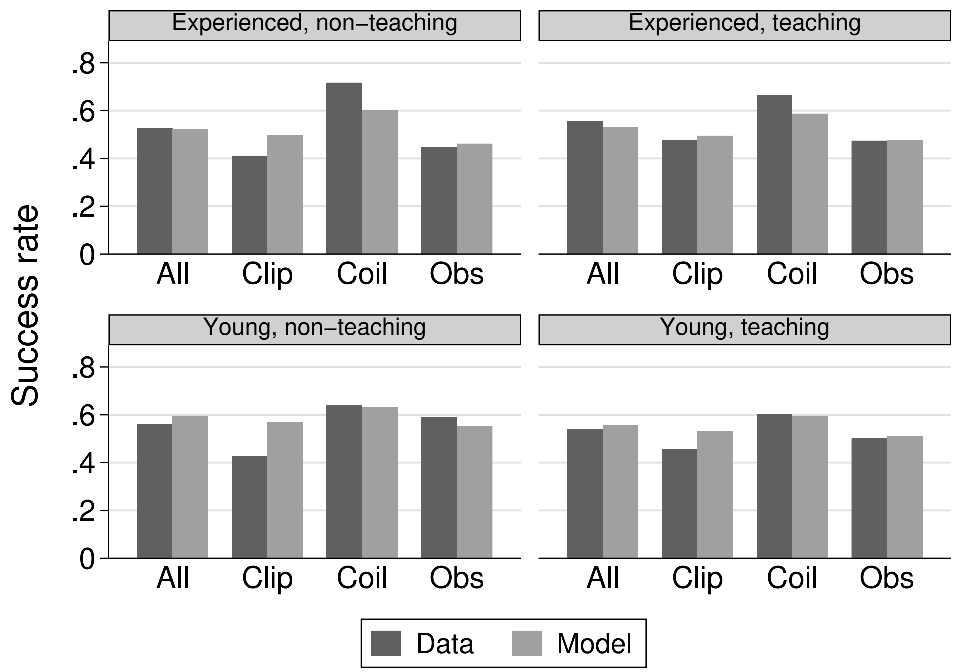


Figure 11: Counterfactual experiment: patient outcomes under uniform payments for clipping and coiling

Notes: The *Data* bars plot the observed patient outcomes by physician groups. The *Model* bars plot the simulated outcomes, assuming uniform payments between clipping and coiling, conditional on patient types and hospitals. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

ficiaries at participating hospitals, and they hope to expand it to all physician providers in 2018. The VM program adjusts the payments on a claim basis for services in the Physician Fee Schedule (PFS). The maximum reward for good patient outcomes is 2% of the baseline fee, and the maximum punishment for adverse outcomes is also 2%. Hence the value at risk for physicians amounts to 4% of the baseline fee.

I explore the impacts of VM-style payment reforms on physician choices and patient outcomes in the context of learning. I assume a simplified payment schedule that pays physicians 102% of the prevailing revenue for a successful treatment and 98% otherwise. Table 18 summarizes the physicians' responses.

	Choice probability			Success	Revenue/case
	Pr(clip)	Pr(coil)	Pr(obs)	rates	($1,000)
			Overall		
Data	0.303	0.409	0.288	0.544	114.25
Model	0.478	0.349	0.173	0.526	159.24
			By physician subgroups		
Experienced physicians at non-teaching hospitals					
Data	0.307	0.341	0.353	0.528	87.96
Model	0.571	0.146	0.283	0.511	130.90
Experienced physicians at teaching hospitals					
Data	0.293	0.431	0.277	0.557	119.00
Model	0.417	0.484	0.100	0.515	180.70
Young physicians at non-teaching hospitals					
Data	0.334	0.460	0.206	0.559	131.16
Model	0.410	0.500	0.090	0.562	161.52
Young physicians at teaching hospitals					
Data	0.242	0.498	0.260	0.541	165.58
Model	0.427	0.448	0.124	0.531	212.56

Table 18: Counterfactual experiment: outcome-contingent payments

NOTES: The *Model* rows report the simulated choice probabilities and treatment success rates, assuming outcome-based payments: physicians receive 102% or 98% of the prevailing revenue when the outcome is a success or a failure, respectively. The results are based on 1,000 simulations using the estimated parameters and the observed sequences of patients. Revenue per case is thousand real 2014 dollars. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

On average, physicians intervene more often under the outcome-based payments. In particular, they tend to choose more clipping but less coiling. The resulting patient outcome, however, is inferior on average. The change seems counterintuitive, but it actually highlights the classic exploitation vs. exploration tradeoff in learning. With payoffs being higher for a success and lower for a failure, physicians would be induced to lean toward the treatment that maximizes current-period outcome. But on the other hand, the outcome-based payments put more emphasis on the future values of learning, which could encourage more experimentation. The overall effects are heterogeneous across physicians and depend on the relative magnitudes of the two opposing forces.

The remainder of Table 18 reflects the heterogeneity across physician groups. The effect on future values of learning dominates for physicians who are experienced and working at teaching hospitals and for those who are young and working at non-teaching hospitals. These two groups of physicians would adopt more coiling, while the other two groups would reduce coiling take-up under the outcome-based payment schedule.

Figures 12 and 13 summarize the responses in treatment choices and patient outcomes by physician groups and by the patients' observed treatment in the data. The two groups of physicians (experienced at teaching hospitals, young at non-teaching hospitals) for whom the future effect dominates tend to choose coiling more often. As for patient outcomes, those who are observed to receive coiling would get worse outcomes under VM-style payments. On the contrary, those who are observed to receive clipping would get moderately better outcomes. This differential response corroborates the above-noted explanation further. The value of learning about clipping, the traditional treatment, is generally lower. Therefore it is more likely that the effect of outcome-based payments on future values is *dominated* by that on the current-period outcome.

7 Concluding Remarks

In this paper, I use the treatment of brain aneurysms as an example to study how two kinds of learning jointly shape physicians' treatment choices: *Bayesian learning* that updates physician beliefs regarding treatment-patient type match value and *learning by doing* that improves treatment-specific surgical skills and reduces the future costs of delivering the same treatment. Using the detailed New York Statewide Inpatient Database (SID) from 2003 to 2014, I retrieve the uninterrupted history of physicians' treatments and outcomes. I then use the novel dataset to uncover empirical patterns that show that both kinds of physician learning are present, and that physicians are forward-looking in their treatment choices.

I develop a dynamic structural model in light of the reduced-form evidence. The model features forward-looking physicians who make treatment choices for heterogeneous patients. My model solution extends the Gittins index policy Gittins (1979) to accommodate a high-dimensional state

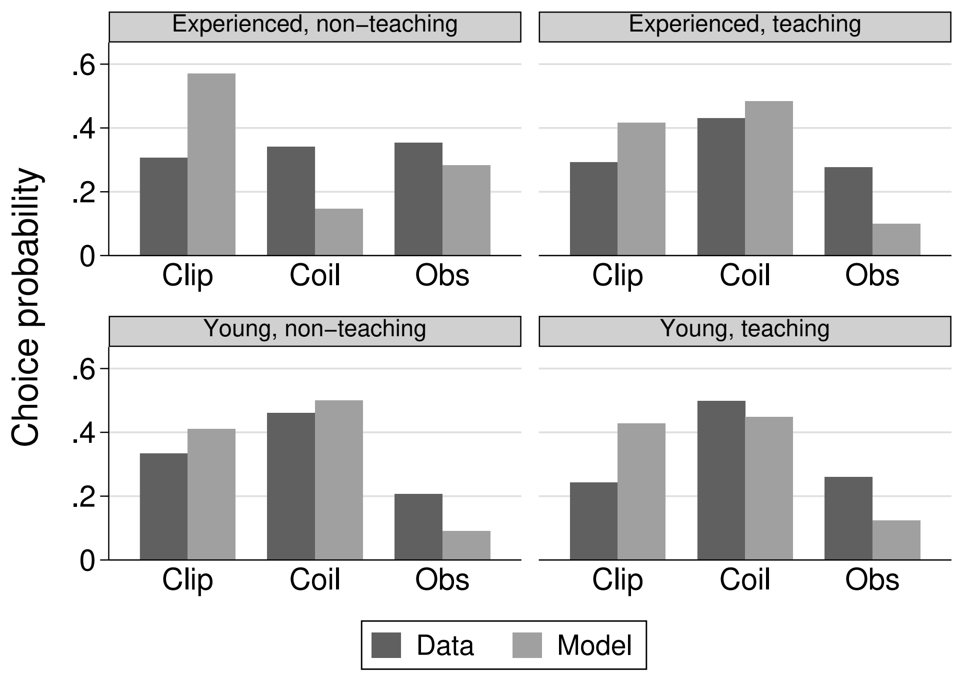


Figure 12: Counterfactual experiment: choice probabilities under outcome-contingent payments
Notes: The *Data* bars plot the observed choice probabilities by physician groups. The *Model* bars plot the simulated probabilities, assuming outcome-based payments: physicians receive 102% or 98% of the prevailing revenue depending on whether the treatment is a success or a failure, respectively.

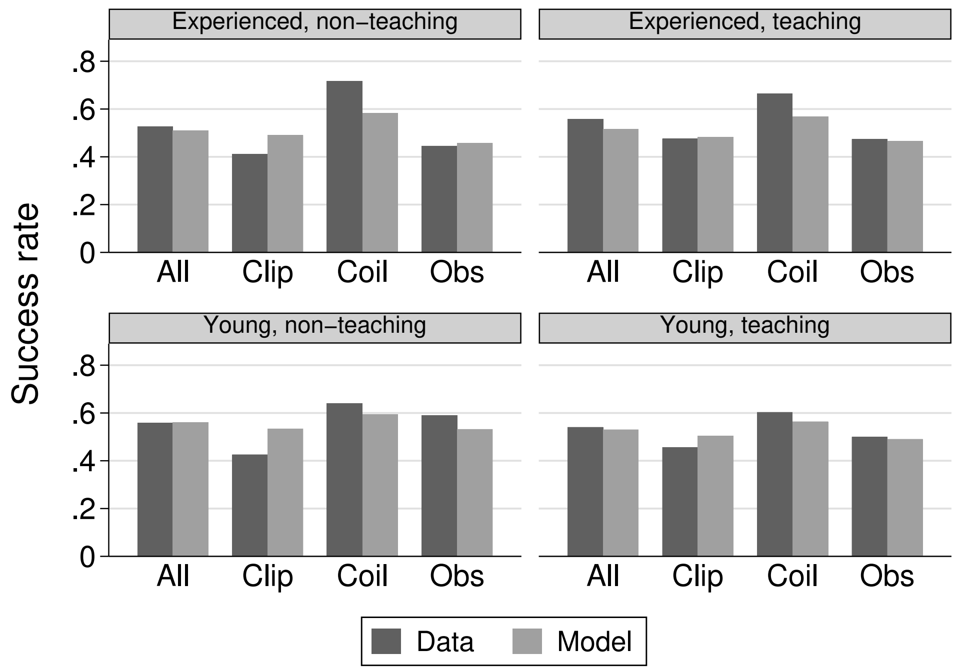


Figure 13: Counterfactual experiment: patient outcomes under outcome-contingent payments
NOTES: The *Data* bars plot the observed treatment success rates by physician groups. The *Model* bars plot the simulated success rates assuming outcome-based payments: physicians receive 102% or 98% of the prevailing revenue when the outcome is a success or a failure, respectively. The results are based on 1,000 simulations using the estimated parameters and the observed sequences of patients. Revenue per case is thousand real 2014 dollars. A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

space and spillover effects between the two kinds of learning. I estimate the model on the SID data and achieve good model fit with the observed choice dynamics. I further disentangle the effects of Bayesian learning and learning by doing: the former dominates the decisions to utilize clipping, the old treatment; the latter dominates the decisions to adopt coiling, the new treatment. Relative magnitudes of the two kinds of learning also vary across physicians. Bayesian learning has more influence on the treatment choices of physicians at non-teaching hospitals, whose prior beliefs about coiling are less optimistic. Learning by doing, on the other hand, is more influential on the decisions of younger physicians, who have lower stocks of surgical skills.

I use the model estimates to compare the observed choices by forward-looking physicians with counterfactual ones by myopic physicians. I find that forward-looking physicians deviate substantially from the myopic best choices, especially on unhealthy patients. I also show that the learning effects are substantial and explain 20% of total variation in the physicians' choice of care, or about one-third of the variation coming from the supply side. Finally, I assess the impacts of alternative payment schedules such as the CMS Value Modifier program that links payments with outcomes. The heterogeneous responses in physician choices highlight the key tradeoff in learning, namely exploiting high flow payoffs versus experimenting with lesser-known options.

This paper sheds light on how belief updating and skill accumulation jointly shape medical decision-making. Moreover, I demonstrate how forward-induction approaches such as the Gittins index can help with the estimation of otherwise high-dimensional models. I also provide new empirical evidence and identify structural learning parameters with the rich variation in the SID data.

My focus in this paper is on individual physicians who make decisions independently and learn from their own experiences. To this end, the treatment of brain aneurysms is particularly suitable because of the scarcity of neurosurgeons and the lack of industry guidelines. Collective or centralized decision-making could be more relevant, however, if a hospital has multiple physicians specializing in the same condition. The hospital may have the dual goal of optimizing patient outcomes *and* the portfolio of its physicians' skills. This is an interesting dimension to explore in future work.

References

Abaluck, Jason, Jonathan Gruber, and Ashley Swanson, "Prescription Drug Use under Medicare Part D: A Linear Model of Nonlinear Budget Sets," *NBER Working Paper*, 2015, *w20976.*

__ **, Leila Agha, Chris Kabrhel, Ali Raja, and Arjun Venkatesh**, "The Determinants of Productivity in Medical Testing: Intensity and Allocation of Care," *American Economic Review*, 2016, *106* (12), 3730–64.

Argote, Linda and Dennis Epple, "Learning Curve in Manufacturing," *Science*, February 1990, *247*, 920–924.

Baicker, Katherine, Elliott S. Fisher, and Amitabh Chandra, "Malpractice Liability Costs And The Practice Of Medicine In The Medicare Program," *Health Affairs*, 2007, *26* (3), 841–852.

Baker, Laurence C., M. Kate Bundorf, and Daniel P. Kessler, "Patients' Preferences Explain A Small But Significant Share Of Regional Variation In Medicare Spending," *Health Affairs*, 2014, *33* (6), 957–963.

Brezzi, Monica and Tze Leung Lai, "Optimal Learning and Experimentation in Bandit Problems," *Journal of Economic Dynamics and Control*, 2002, *27* (1), 87–108.

Brinjikji, W., A.A. Rabinstein, D.M. Nasr, G. Lanzino, D.F. Kallmes, and H.J. Cloft, "Better Outcomes with Treatment by Coiling Relative to Clipping of Unruptured Intracranial Aneurysms in the United States, 2001–2008," *American Journal of Neuroradiology*, 2011, *32* (6), 1071–1075.

Chan, David C., "Information Frictions and Practice Variation: Evidence from Physicians in Training," *Working Paper, Standford University*, 2016.

Chandra, Amitabh and Douglas O Staiger, "Productivity Spillovers in Healthcare: Evidence from the Treatment of Heart Attacks," *The Journal of Political Economy*, 2007, *115*, 103.

__ **, David Cutler, and Zirui Song**, "Chapter Six - Who Ordered That? The Economics of Treatment Choices in Medical Care," in Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros, eds., *Handbook of Health Economics*, Vol. 2, Elsevier, 2011, pp. 397 – 432.

Ching, Andrew T, Tülin Erdem, and Michael P Keane, "Learning Models: An Assessment of Progress, Challenges, and New Developments," *Marketing Science*, 2013, *32* (6), 913–938.

Coscelli, Andrea and Matthew Shum, "An Empirical Model of Learning and Patient Spillovers in New Drug Entry," *Journal of Econometrics*, 2004, *122* (2), 213–246.

Crawford, Gregory S and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (4), 1137–1173.

Currie, Janet and W Bentley MacLeod, "Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians," *Journal of Labor Economics*, 2017, *35* (1-43).

_, W. Bentley MacLeod, and Jessica Van Parys, "Provider Practice Style and Patient Health Outcomes: The Case of Heart Attacks," *Journal of Health Economics*, 2016, *47*, 64 – 80.

Cutler, David, Jonathan Skinner, Ariel Dora Stern, and David Wennberg, "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending," *NBER Working Paper*, 2013, *w19320.*

Dickstein, Michael J, "Efficient Provision of Experience Goods: Evidence from Antidepressant Choice," *Standford University Working Paper*, 2014.

Ferreyra, Maria Marta and Grigory Kosenok, "Learning about New Products: An Empirical Study of Physicians' Behavior," *Economic Inquiry*, 2011, *49* (3), 876–898.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Health Care: Evidence from Patient Migration," *The Quarterly Journal of Economics*, 2016, *131* (4), 1681–1726.

Gittins, John C, "Bandit Processes and Dynamic Allocation Indices," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1979, *41* (2), 148–177.

Gittins, John, Kevin Glazebrook, and Richard Weber, *Multi-Armed Bandit Allocation Indices*, John Wiley & Sons, Ltd., 2011.

Gruber, Jon, John Kim, and Dina Mayzlin, "Physician Fees and Procedure Intensity: the Case of Cesarean Delivery," *Journal of Health Economics*, 1999, *18* (4), 473–490.

Ho, Vivian, "Learning and the Evolution of Medical Technologies: the Diffusion of Coronary Angioplasty," *Journal of Health Economics*, 2002, *21* (2002), 873–885.

Johnson, Erin M and M Marit Rehavi, "Physicians Treating Physicians: Information and Incentives in Childbirth," *American Economic Journal: Economic Policy*, 2016, *8* (1), 115–141.

Jovanovic, Boyan, "Job Matching and the Theory of Turnover," *Journal of Political*, 1979, *87* (5), 972–990.

_ **and Yaw Nyarko**, "Stepping-Stone Mobility," *Carnegie-Rochester Conference Series on Public Policy*, 1997, *46*, 289–325.

Kolstad, Jonathan T., "Information and Quality When Motivation Is Intrinsic: Evidence from Surgeon Report Cards," *American Economic Review*, December 2013, *103* (7), 2875–2910.

Miller, Robert A., "Job Matching and Occupational Choice," *Journal of Political Economy*, 1984, *92* (6), 1086–1120.

Molitor, David, "The Evolution of Physician Practice Style: Evidence from Cardiologist Migration," *American Economic Journal: Economic Policy*, 2018, *Forthcoming*.

Narayanan, Sridhar and Puneet Manchanda, "Heterogeneous Learning and the Targeting of Marketing Communication for New Products," *Marketing Science*, 2009, *28* (3), 424–441.

Papageorgiou, Theodore, "Learning Your Comparative Advantages," *Review of Economic Studies*, 2014, *81*, 1263–1295.

Reagans, Ray, Linda Argote, and Daria Brooks, "Individual Experience and Experience Working Together: Predicting Learning Rates from Knowing Who Knows What and Knowing How to Work Together," *Management Science*, June 2005, *51* (6), 869–881.

Saxell, Tanja, "Private Experience and Observational Learning in Pharmaceutical Demand," *Working Paper, Government Institute for Economic Research, Finland*, 2014.

Skinner, Jonathan, "Chapter Two - Causes and Consequences of Regional Variations in Health Care," in Mark V. Pauly, Thomas G. Mcguire, and Pedro P. Barros, eds., *Handbook of Health Economics*, Vol. 2, Elsevier, 2011, pp. 45 – 93.

Stone, Scellig and Mark Bernstein, "Prospective Error Recording in Surgery: An Analysis of 1108 Elective Neurosurgical Cases," *Neurosurgery*, 2007, *60* (6), 1075–1082.

The Dartmouth Atlas Working Group, *Variation in the Care of Surgical Conditions*, The Dartmouth Institute for Health Policy & Clinical Practice, 2015.

Vlak, Monique HM, Ale Algra, Raya Brandenburg, and Gabriël JE Rinkel, "Prevalence of Unruptured Intracranial Aneurysms, with Emphasis on Sex, Age, Comorbidity, Country, and Time Period: A Systematic Review and Meta-Analysis," *The Lancet Neurology*, 2011, *10* (7), 626 – 636.

Whittle, Peter, "Multi-armed Bandits and the Gittins Index," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1980, *42* (2), 143–149.

_ , *Optimization Over Time*, Vol. 1, John Wiley, 1982.

_ , "Restless Bandits: Activity Allocation in a Changing World," *Journal of Applied Probability*, 1988, *25* (A), 287–298.

Wiebers, David O, International Study of Unruptured Intracranial Aneurysms Investigators et al., "Unruptured Intracranial Aneurysms: Natural History, Clinical Outcome, and Risks of Surgical and Endovascular Treatment," *The Lancet*, 2003, *362* (9378), 103–110.

Zacharia, Brad E., Samuel S. Bruce, Amanda M. Carpenter, Zachary L. Hickman, Kerry A. Vaughan, Catherine Richards, William E. Gold, June Lu, Geoffrey Appelboom, Robert A. Solomon, and E. Sander Connolly, "Variability in Outcome After Elective Cerebral Aneurysm Repair in High-Volume Academic Medical Centers," *Stroke*, 2014, *45* (5), 1447–1452.

A Proof of Proposition 1

A.1 Overview of the proof

In this section, I present the complete proof of Proposition 1. First, recall the definition of the Gittins index (same as Definition 1 and Proposition 1 in the main text):

Definition 1. *For arm d in state $(\boldsymbol{\theta}^d, e^d, k, \boldsymbol{r}^d)$, construct a two-armed bandit process by adding an auxiliary arm with lump-sum retirement payment M. The Gittins index for arm d, $M^{dk}(\boldsymbol{\theta}^d, e^d, \boldsymbol{r}^d)$, is the infimum of all the M values that the physician is willing to take and retire. That is*

$$M^{dk}(\boldsymbol{\theta}^d, e^d, \boldsymbol{r}^d) := \inf_M \{M : \ \phi^d(\boldsymbol{\theta}^d, e^d, k, \boldsymbol{r}^d, M) = M\} \tag{A1}$$

Proposition 1. *Under Assumption 1, the modified Gittins index policy that always selects the treatment option with the highest $M^{dk}(\boldsymbol{\theta}^d, e^d, \boldsymbol{r}^d)$ is optimal for the physician's problem (10):*

$$\max_{\{d_t\}_{t=0,1,\dots}} \mathbb{E}[\sum_{t=0}^{\infty} \beta_i^t u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk}) \mid \boldsymbol{\theta}_0, \boldsymbol{e}_0, k_0, r_0^{dk}] \tag{A2}$$

I follow the approach of Whittle (1982) and prove the proposition in 4 steps: (1) I start by defining the one-step forward operator, L^d, and show that the optimal payoff from A2, G, must also be optimal for the sequential one-step forward problem; (2) I then examine the optimal payoff of the modified MAB, Φ, and establish the properties Φ must satisfy; (3) I construct a candidate payoff, $\hat{\Phi}$, whose value is determined jointly by indexes of all arms given the realized patient type k, $\{M^{dk}\}_{d\in D}$; (4) finally, I show that $\hat{\Phi}$ is indeed the optimal payoff, Φ, and the optimality is achieved by an index policy that always prescribes the arm with the highest M^{dk}. Hence I have shown that the Gittins index policy is optimal for the physician's problem.

A.2 Step 1: Optimal payoff and the one-step forward problem

Let constants A and B be the uniform lower and upper bounds of *total* discounted payoffs, respectively:

$$-\infty < A(1-\beta) \leq u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk}) \leq B(1-\beta) < \infty \tag{A3}$$

Denote the maximum expected total discounted payoff function for the physician's problem (10) by G

$$G(\boldsymbol{\theta}_0, \boldsymbol{e}_0, k_0, \boldsymbol{r}_0) = \max_{\pi} \mathbb{E}_{\pi}[\sum_{t=1}^{\infty} \beta^t u_t \mid \boldsymbol{\theta}_0, \boldsymbol{e}_0, k_0, \boldsymbol{r}_0] \tag{A4}$$

where π is a policy and $(\boldsymbol{\theta}_0, \boldsymbol{e}_0, k_0)$ are the initial states. Then G must satisfy

$$G = \max_{d} L^d G \tag{A5}$$

with the one-step operator L^d defined as

$$L^d G(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}) = u^d(\boldsymbol{\theta}^d, e^d, k, r^{dk}) + \beta \mathbb{E}[G(\boldsymbol{\theta}', \boldsymbol{e}', k', \boldsymbol{r}') \mid d, \boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}] \tag{A6}$$

The one-step transition of states is such that $\boldsymbol{\theta}'$ and $\boldsymbol{\theta}$ only differs in θ^{dk}; $\boldsymbol{e}'$ and $\boldsymbol{e}$ only differs in that $e^{d\prime} = e^d + 1$; k' is a new draw from the same distribution as k_t; and $\boldsymbol{r}'$ is expected by the physician to remain unchanged.

A.3 Step 2: Modified MAB with an auxiliary arm for retirement

Now consider a modified version of the physician's problem with the added option of taking a given lump-sum payment M and retiring. Define the maximum expected total discounted payoff function for this modified problem by Φ. Then we have

$$\Phi(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M) = \max \left\{ M, \max_{d} L^d \Phi(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}) \right\} \tag{A7}$$

Let τ be the period in which the retirement option is chosen, then

$$\Phi = V_c + M \mathbb{E}[\beta^{\tau} \mid \boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}] \tag{A8}$$

where V_c is the total discounted payoff from running the MAB before retirement. Note that $\tau = +\infty$ is allowed, in which case retirement is never chosen and $\mathbb{E}[\beta^{\tau}] = 0$.

Similarly, define a modified two-armed bandit process with one arm being option d and the other being retirement with lump-sum payment M. Let ϕ^d be the maximal expected payoff function, which must solve

$$\begin{aligned}\phi^d &= \max\left\{M, L^d\phi^d\right\} && \text{(A9)}\\ &= \max\left\{M, u^d(\boldsymbol{\theta}^d, e^d, k, r^{dk}) + \beta\mathbb{E}[\phi(\boldsymbol{\theta^{d\prime}}, e^{d\prime}, k', \boldsymbol{r'}) \mid d, \boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}]\right\} && \text{(A10)}\end{aligned}$$

The following properties of Φ and ϕ^d will become useful in the proof of the optimality of an index policy.

Lemma A1. $\Phi(\boldsymbol{\theta}, \boldsymbol{e}, k, M)$ *is non-decreasing, convex in* M. *And*

$$\Phi(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M) = \begin{cases} G(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}) & \textit{for } M \leq A \\ M & \textit{for } M \geq B \end{cases} \tag{A11}$$

Proof. That Φ is non-decreasing in M and relationship (A11) are obvious: increasing the payment for retirement cannot make the physician worse off; the retirement option is disregarded when M is smaller than the minimum total discounted payoff A, and it is taken immediately ($\tau = 0$) when M exceeds the maximal total discounted payoff B. The convexity of Φ in M follows from (A11) and (A8). □

A.4 Step 3: A candidate for Φ

Define $\hat{\Phi}$ as a candidate for the optimal expected payoff function

$$\hat{\Phi}(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M) = B - \int_M^B \prod_d \frac{\partial\phi_d(\boldsymbol{\theta}^d, e^d, k, r^{dk}, m)}{\partial m} \mathrm{d}m \tag{A12}$$

Integration by parts yields (suppressing arguments for $\hat{\Phi}$)

$$\hat{\Phi} = \phi^d(\boldsymbol{\theta}^d, e^d, k, r^{dk}, M)P^d(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M) + \int_M^\infty \phi^d(\boldsymbol{\theta}^d, e^d, k, r^{dk}, m)\mathrm{d}_m P^d(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, m) \tag{A13}$$

where P^d is defined as

$$P^d(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M) := \prod_{d' \neq d} \frac{\partial\phi_{d'}(\boldsymbol{\theta}^{d'}, e^{d'}, k, r^{d'k}, M)}{\partial M} \tag{A14}$$

Lemma A2. $P^d(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M)$ *is non-negative and non-decreasing in* M*; and is equal to 1 when*

$$M > \max_{d' \neq d} M^{d'k} \tag{A15}$$

Proof. By Lemma A1, ϕ^d is also non-decreasing and convex in M. Hence $\partial\phi_d/\partial M$ is non-negative and non-decreasing in M. It also follows from Lemma A1 that $\phi^d = M$ when $M \geq M^{dk}$, where M^{dk} is arm d's Gittins index in Definition (1). Thus $P^d = 1$ when M is no smaller than the index for any arm $d' \neq d$. □

Before proving the optimality of $\hat{\Phi}$, note that $\hat{\Phi}$ is constructed based on the following observation:

$$\frac{\partial\Phi(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M)}{\partial M} = \mathbb{E}[\beta^\tau \mid \boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}] = \prod_{d'} \mathbb{E}[\beta^{\tau_{d'}} \mid \boldsymbol{\theta}^{d'}, e^{d'}, k, r^{d'k}] = \prod_{d'} \frac{\partial\phi_{d'}(\boldsymbol{\theta}^{d'}, e^{d'}, k, r^{d'k}, M)}{\partial M} \tag{A16}$$

where the first equality follows (A8) and the second equality from the independence across arms.[26]

A.5 Step 4: Optimality of $\hat{\Phi}$ and the policy that achieves the optimality

To show that $\hat{\Phi}$ is indeed the optimal payoff under a Gittins index policy, it suffices to show that it solves Equation (A7) by always choosing the arm with the largest index, M^{dk}.

I begin by showing the first part of (A7), i.e. $\hat{\Phi} \geq M$:

Lemma A3. $\hat{\Phi} \geq M$*, with equality when*

$$M > \max_{d' \in \mathcal{D}} M^{d'k} \tag{A17}$$

Proof. Suppress the state variables $(\boldsymbol{\theta}, \boldsymbol{e}, k)$ for simplicity of exposition in this proof. Recall that $\partial\phi/\partial M$ is at most 1, hence

$$\begin{aligned} \hat{\Phi}(M) &= B - \int_M^B \prod_{d'} \frac{\partial\phi_{d'}(m)}{\partial m} \mathrm{d}m && \text{(A18)} \\ &\geq B - \int_M^B 1 \, \mathrm{d}m && \text{(A19)} \\ &= M && \text{(A20)} \end{aligned}$$

And the weak inequality (A20) becomes equality only when $\partial\phi_{d'}(m)/\partial m = 1$ for all $d' \in \mathcal{D}$. This

[26] Whittle (1980) observed that P^d (and $\partial\phi^d/\partial m$) is essentially a distribution function.

coincides with condition (A17) by definition of the Gittins index $M_{d'}$. $\square$

Now I show that $\hat{\Phi}$ also satisfies the second half of (A7):

Lemma A4. *$\hat{\Phi} \geq L^d \hat{\Phi}$, with equality when*

$$M^{dk} = \max_{d' \in \mathcal{D}} M^{d'k} \geq M \tag{A21}$$

Proof. First consider the effect of applying the one-step operator, L^d on $P^d(\boldsymbol{\theta}, \boldsymbol{e}, k, M)$:

$$\begin{aligned}
L^d P^d(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M) &:= L^d \left[\prod_{d' \neq d} \frac{\partial \phi_{d'}(\boldsymbol{\theta}^{d'}, e^{d'}, k, r^{d'k}, M)}{\partial M} \right] && \text{(A22)}\\
&= L^d \left[\prod_{d' \neq d} \mathbb{E}[\beta^{\tau_{d'}} \mid \boldsymbol{\theta}^{d'}, e^{d'}, k, r^{d'k}] \right] && \text{(A23)}\\
&= L^d \left[\prod_{d' \neq d} \mathbb{E}[\beta^{\tau_{d'}} \mid \boldsymbol{\theta}^{d'}, e^{d'}, r^{d'k}] \right] && \text{(A24)}\\
&= \prod_{d' \neq d} \mathbb{E}[\beta^{\tau_{d'}} \mid \boldsymbol{\theta}^{d'}, e^{d'}, r^{d'k}] && \text{(A25)}\\
&= \prod_{d' \neq d} \mathbb{E}[\beta^{\tau_{d'}} \mid \boldsymbol{\theta}^{d'}, e^{d'}, k, r^{d'k}] && \text{(A26)}\\
&= P^d(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M) && \text{(A27)}
\end{aligned}$$

Equality (A24) holds by Assumption (1), which abstracts away potentially persistent effects of the current patient type k on the duration of trial, $\tau_{d'}$. And Equality (A26) holds because the state variables $(\boldsymbol{\theta}^{d'}, e^{d'},)$ remain frozen when the one-step operator L^d chooses to operate arm $d \neq d'$.

Now apply L^d to $\hat{\Phi}$ and we get

$$L^d \hat{\Phi} = L^d \phi^d(\boldsymbol{\theta}^d, \boldsymbol{e}^d, k, \boldsymbol{r}, M) P^d(\boldsymbol{\theta}, \boldsymbol{e}, k, \boldsymbol{r}, M) + \int_M^\infty L^d \phi^d(\boldsymbol{\theta}^d, \boldsymbol{e}^d, k, \boldsymbol{r}, m) \mathrm{d}_m P^d(\boldsymbol{\theta}, \boldsymbol{e}, \boldsymbol{r}, k, m) \tag{A28}$$

Suppressing the state variables, we get the following relationship between $\hat{\Phi}$ and $L^d \hat{\Phi}$:

$$\hat{\Phi} - L^d \hat{\Phi} = \left[\phi^d(M) - L^d \phi^d(M) \right] P^d(M) + \int_M^\infty \left[\phi^d(m) - L^d \phi^d(m) \right] \mathrm{d}_m P^d(m) \tag{A29}$$

Because ϕ solves the problem (A10), we have $\left[\phi^d(M) - L^d \phi^d(M) \right] \geq 0$. Moreover, P^d is non-negative. Hence $\hat{\Phi} \geq L^d \hat{\Phi}$.

Moreover, we know that $\phi^d(M) = L^d \phi^d(M)$ when $M \leq M^{dk}$. And that P^d is non-decreasing in

M, and is equal to 1 when $M > \max\limits_{d' \neq d} M^{d'k}$. These two restrictions on M jointly implies that when $M^{dk} = \max\limits_{d' \neq d} M^{d'k} \geq M$, we have $\hat{\Phi} = L^d \hat{\Phi}$. $\square$

It is now easy to prove Proposition 1. Lemmas A3 and A4 show that the candidate solution $\hat{\Phi}$ indeed solves (A10), $\Phi = \max\left\{M, \max_d L^d \Phi\right\}$. Furthermore, the conditions on M in Lemmas A3 and A4 prescribe a policy for achieving the maximal expected payoff, which is to always choose d for which $M^{dk} = \max\limits_{d' \neq d} M^{d'k} \geq M$, i.e. the arm with the highest Gittins index.

B Numerical Approximation of the Modified Gittins index

B.1 Overview of the method

In this section, I derive the approximate Gittins index (13) and its full specification. I adapt the approximation by Brezzi and Lai (2002), noting that their work is based on Gittins' formulation of the MAB process, which differs from the formulation I use in the paper following Whittle (1982). The two approaches differ in the specification of the auxiliary arm: Gittins assumes the agent receives a fixed rate v in each period she chooses the auxiliary arm, and that she still has the option of choosing the original arm, d; Whittle assumes that the agent receives a one-time, lump-sum retirement pay M when she choses the auxiliary arm, and that she can never operate the two-armed bandit machine again. They use v and M as indexes, respectively.

Whittle (1982) compared the two approaches, noting that Gittins' formulation gives rise to an intuitive optimal stopping argument: between the auxiliary arm that pays a fixed rate v and a risky arm, d, if the agent chooses former in some period, it must be that she values v more than the combined value of d's flow and continuation payoff. Further, the states of d will remain frozen when it is inactive; the agent will then keep choosing the same v in the next period, and for that matter all subsequent periods, hence the optimal stopping rule. Whittle observes that Gittins' formulation of the auxiliary arm elucidates the economic intuition of the index, but makes the proof of optimality hard to follow (Whittle, 1980). Whittle's own formulation of the auxiliary arm does not employ an optimal stopping rule–choosing the arm directly stops all future decision-making, but renders a simple proof.

Beyond these methodological distinctions, indexes under the two approaches do not differ fundamentally, and they are linked by the following relationship:

$$M = \frac{v}{1 - \beta} \tag{A30}$$

where β is the discount factor (Whittle, 1982).

I draw on the approximation of the Gittins index by Brezzi and Lai (2002) and adapt it to my model in three ways: First, I convert Brezzi and Lai's approximation (derived under Gittins' formulation) to one that suits the Whittle's formulation that I use. Second, I handle the uncertain starting states (in period $(t+1)$) of the index in my model. Finally, I complete the approximated Gittins index, which only accounts for Bayesian learning, by adding the learning-by-doing effects and the financial incentives.

B.2 The flow utility

The complete index for my model starts with the flow utility

$$\text{Index } M_t^{dk}\text{, part 1:} \quad u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk}) = \mu_t^{dk} + \alpha r_t^{dk} - c(e_t^d) \tag{A31}$$

I separate the flow utility from the remainder of the problem for two reasons: first, the effect of the *current* patient's type is assumed to be transitory, thus only showing up in the flow utility; second, as discussed in the paper, the physician virtually has one "average" type of patient in future periods from an *ex ante* perspective. The second point implies that from period $(t+1)$ onwards, the value of each arm can be summarized in a standard Gittins index that no longer depends on patient types.

B.3 Step 1: approximated Gittins index for the standard MAB

From period $(t+1)$ onward, it is as if the physician expects to have only one "average" type of patient. Suppose the physician's posterior belief about d's match value for the average type, after observing period t outcomes, has mean and variance μ_{t+1}^d and ν_{t+1}^d, respectively. Then following Brezzi and Lai (2002), the closed-form approximation to the Gittins index is

$$\text{Index } M_t^{dk}\text{, part 2 (conditional on } y_t^d\text{):} \quad \tilde{M}(\boldsymbol{\theta}_{t+1}^d, e_{t+1}^d) = (1-\beta)^{-1}\left[\mu_{t+1}^d + \sqrt{\nu_{t+1}^d}\,\psi\Big(\frac{\nu_{t+1}^d}{-\ln(\beta)\sigma_{\mu,d}^2}\Big)\right] \tag{A32}$$

where the $(1-\beta)$ in the denominator comes from using (A30) to convert the index under Gittins' formulation to one under Whittle's. $\sigma^2_{\mu,d} = \mu^d_{t+1}(1-\mu^d_{t+1})$ and ψ is the single-argument function:

$$\psi(s) = \begin{cases} \sqrt{s/2}, & s \leq 0.2 \\ 0.49 - 0.11s^{-0.5}, & 0.2 < s \leq 1 \\ 0.63 - 0.26s^{-0.5}, & 1 < s \leq 5 \\ 0.77 - 0.58s^{-0.5}, & 5 < s \leq 15 \\ \sqrt{2\ln s - \ln(\ln s) - \ln(16\pi)}, & s > 15 \end{cases} \tag{A33}$$

The posterior mean and variance, μ^d_{t+1} and ν^d_{t+1}, are intuitive to get. Suppose the physician has observed the period t outcome and updated her belief about d's match value with each type k.[27] Denote the posterior mean and variance as μ^{dk}_{t+1} and ν^{dk}_{t+1}, respectively. The physician simply uses the weighted average of type-specific posterior means and variances to calculate μ^d_{t+1} and ν^d_{t+1}:

$$\mu^d_{t+1} = \sum_{k=1}^{K} \lambda_k \mu^{dk}_{t+1}, \qquad \mu^{dk}_{t+1} = \frac{a^{dk}_{t+1}}{a^{dk}_{t+1} + b^{dk}_{t+1}} \tag{A34}$$

$$\nu^d_{t+1} = \sum_{k=1}^{K} \lambda_k^2 \nu^{dk}_{t+1}, \qquad \nu^{dk}_{t+1} = \frac{a^{dk}_{t+1} b^{dk}_{t+1}}{(a^{dk}_{t+1} + b^{dk}_{t+1})^2 (a^{dk}_{t+1} + b^{dk}_{t+1} + 1)} \tag{A35}$$

where λ_k is the type-specific patient arrival rate; a^{dk}_{t+1} and b^{dk}_{t+1} are the number of previous successes and failures for the (d,k) combination up to, but not including $(t+1)$.

B.4 Step 2: taking expectation over the period-t outcome realization

The posterior means and variances, μ^{dk}_{t+1} and ν^{dk}_{t+1}, are the results of physician belief updating *after* she observes period-t outcomes. The outcome is not yet known to the physician when she is to make the period-t choice. Accordingly, she takes expectation over the two possible realizations of y^{dk}_t, with the probability of $y^{dk}_t = 1$ being μ^{dk}_t:

$$\text{Index } M^{dk}_t\text{, part 2 (expectation over } y^d_t\text{):} \quad \mathbb{E}\left[\tilde{M}(\boldsymbol{\theta}^d_{t+1}, e^d_{t+1})\right] \tag{A36}$$

[27] The beliefs regarding d's match value for types that didn't arrive in period t remain the same at the beginning of period $(t+1)$ as at the beginning of period t.

B.5 Step 3: adjusting for learning-by-doing effects and financial incentives

So far the index has not captured the continuation value from learning by doing and future financial incentives. Suppose the partial index $\tilde{M}$ prescribes a stopping time τ^*. Then one can sum the financial incentives, net of the cost of furnishing treatment, as

$$\text{Index } M_t^{dk}\text{, part 3:} \quad \sum_{\tau=0}^{\infty} \beta^{\tau} \left(\alpha \mathbb{E}(r_{t+1+\tau}^d) - c(e_{t+1+\tau}^d) \right) \tag{A37}$$

where $\mathbb{E}(r_{t+1+\tau}^d) = \sum_{k=1}^{K} \lambda_k r_{t+1+\tau}^{dk}$, is the expected revenue for the "average" type.

B.6 Complete specification

Assembling the parts derived above, I get the following complete specification of the Gittins index, adapted from Brezzi and Lai (2002), to accommodate two types of learning in this paper:

$$\begin{aligned} M^{dk}(\boldsymbol{\theta}_t^d, e_t^d, \boldsymbol{r}_t^d) &= u_t^d(\boldsymbol{\theta}_t^d, e_t^d, k_t, r_t^{dk}) \\ &\quad + \beta \mathbb{E}\left[\tilde{M}(\boldsymbol{\theta}_{t+1}^d, e_{t+1}^d) + \sum_{\tau=0}^{\tau^*} \beta^{\tau} \left(\alpha \mathbb{E}(r_{t+1+\tau}^d) - c(e_{t+1+\tau}^d) \right) \Big| \boldsymbol{\theta}_t^d, e_t^d, k_t \right] \end{aligned} \tag{A38}$$

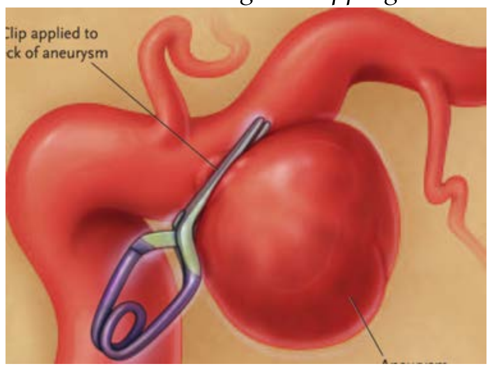
I rearrange the above to get

$$\begin{aligned} & M^{dk}(\boldsymbol{\theta}_t^d, e_t^d, \boldsymbol{r}_t^d) \\ = \; & \mu_t^{dk} + \alpha r_t^{dk} - c(e_t^d) \\ & + \frac{\alpha\beta}{1-\beta} \sum_{k=1}^{K} \lambda_k r_t^{dk} - \frac{\beta \exp(-\alpha_2^c)}{1 - \beta \exp(-\alpha_2^c)} c(e_t^d) \\ & + \frac{\beta}{1-\beta} \left[\mu_t^{dk} \times \left(\mu_{t+1}^{d\oplus} + \sqrt{\nu_{t+1}^{d\oplus}} \psi^{\oplus} \right) + (1 - \mu_t^{dk}) \times \left(\mu_{t+1}^{d\ominus} + \sqrt{\nu_{t+1}^{d\ominus}} \psi^{\ominus} \right) \right] \end{aligned} \tag{A39}$$

$\mu_{t+1}^{d\oplus}$ is the average posterior mean after observing $y_t^d = 1$, which happens with probability μ_t^{dk} as perceived by the physician; and $\mu_{t+1}^{d\ominus}$ is that after observing $y_t^d = 0$, which happens with perceived probability $(1 - \mu_t^{dk})$. $\nu_{t+1}^{d\oplus}, \nu_{t+1}^{d\ominus}, \psi_{t+1}^{d\oplus}$, and $\psi_{t+1}^{d\ominus}$ are defined similarly.

C Supplemental Figures

Panel A. Surgical clipping



Panel B. Endovascular coiling

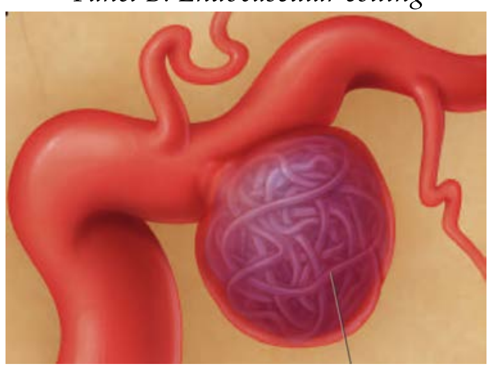


Figure A1: Illustration of clipping and coiling for brain aneurysms

Source: *Brisman JL, Song JK, Newell DW. Cerebral aneurysms. NEJM 2006; 355:928-39.*

Notes: The top figure illustrates the treatment of brain aneurysms with surgical clipping. With craniotomy, the neurosurgeon places the small clip across the neck of the aneurysm to stop the artery blood from flowing into the aneurysm. The bottom figure illustrates the treatment with endovascular coiling. The neurosurgeon inserts a fine platinum coiling into the groin artery and navigates the blood vessel with the help of a catheter. Once the coil reaches the bottom of the aneurysm, it is wound up into a ball that fills the aneurysm, thereby reducing the inflow of blood.

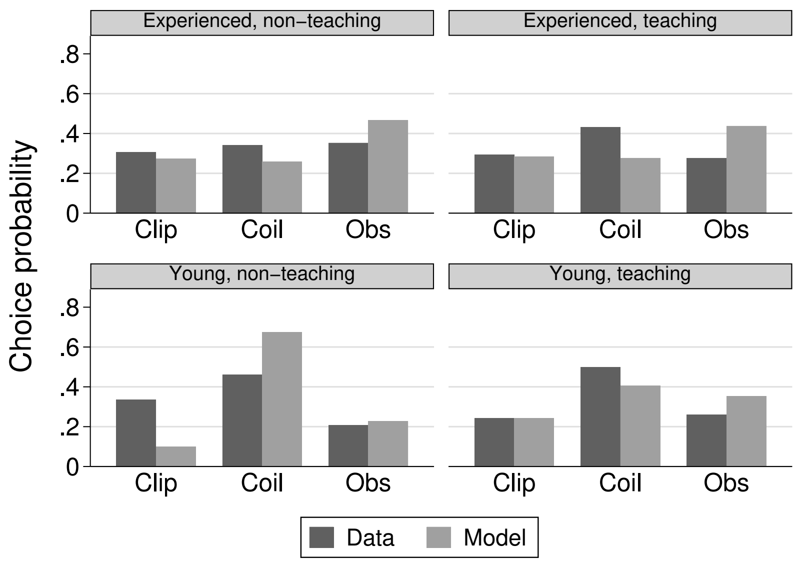


Figure A2: Counterfactual experiment: choice probabilities when physicians maximize flow patient outcome (by physician characteristics)

Notes: The *Data* bars plot the observed choice probabilities. The *Model* bars plot the simulated choice probabilities and patient outcomes, assuming physicians maximize current-period expected patient outcomes (plus a treatment-specific, Type-I extreme value error). *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

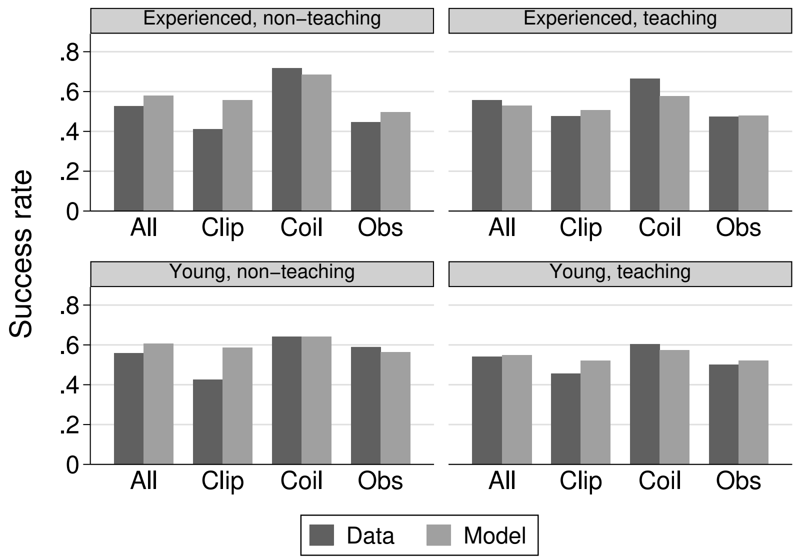


Figure A3: Counterfactual experiment: patient outcomes when physicians maximize flow patient outcome (by physician characteristics)

NOTES: The *Data* bars plot the observed patient outcomes, i.e. the treatment success rates. The *Model* bars plot the simulated success rates, assuming physicians maximize current-period expected patient outcomes (plus a treatment-specific, Type-I extreme value error). A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

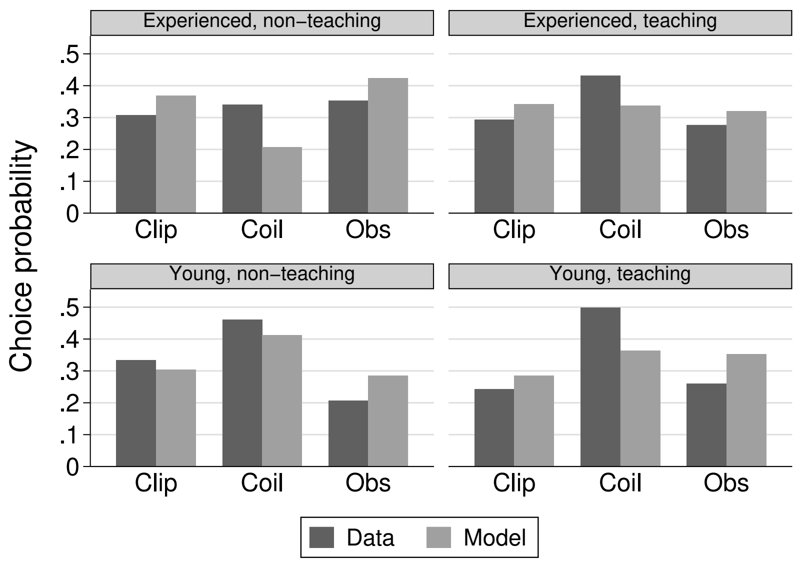


Figure A4: Counterfactual experiment: choice probabilities when physicians maximize flow payoff (by physician characteristics)

Notes: The *Data* bars plot the observed choice probabilities. The *Model* bars plot the simulated choice probabilities, assuming physicians maximize their own flow payoffs (plus a treatment-specific, Type-I extreme value error). *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.

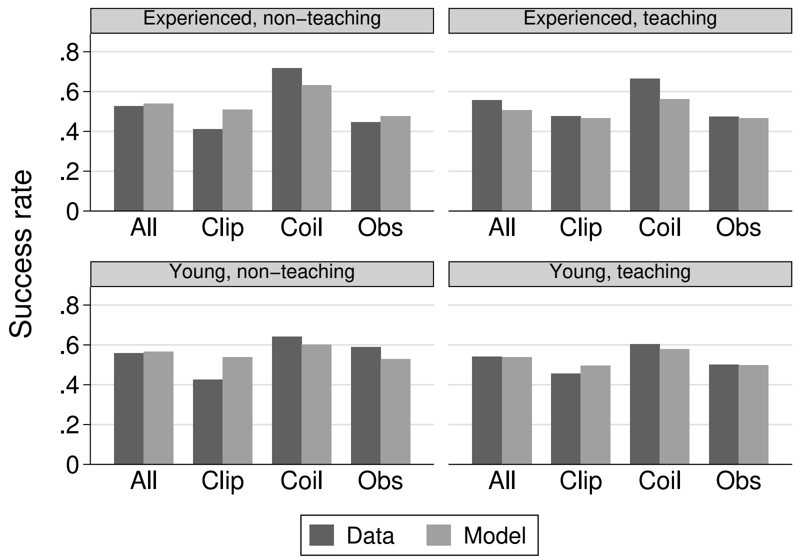


Figure A5: Counterfactual experiment: patient outcomes when physicians maximize flow payoff (by physician characteristics)

NOTES: The *Data* bars plot the observed treatment success rates. The *Model* bars plot the simulated success rates, assuming physicians maximize their own flow payoffs (plus a treatment-specific, Type-I extreme value error). A treatment is a success if the patient can be discharged home and does not need assisted care. *Clip* stands for surgical clipping; *coil* stands for endovascular coiling; *obs* stands for watchful observation, i.e. no intervention. Young physicians are those with no more than 5 years of experience in the data. Teaching-hospital physicians are those working primarily (measured by caseload) at a teaching hospital. If a physician works at multiple hospitals, the one with the largest share of the physician's cases is used.